(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
30 November 2000 (30.11.2000)

PCT

(10) International Publication Number
WO 00/71578 A2

(51) International Patent Classification⁷: C07K 14/15

(21) International Application Number: PCT/EP00/04534

(22) International Filing Date: 19 May 2000 (19.05.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
99401220.1 20 May 1999 (20.05.1999) EP

(71) Applicant *(for all designated States except US)*: **CNRS CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE** [FR/FR]; 3, rue Michel-Ange, F-75794 Paris Cedex 16 (FR).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **COSSET, François-Loïc** [FR/FR]; 30, Grande Rue de la Guillotière, F-69007 Lyon (FR). **LAVILLETTE, Dimitri** [FR/FR]; 32 bis, rue du Lieutenant Colonel Girard, F-69007 Lyon (FR).

(74) Agents: GROSSET-FOURNIER, Chantal et al.; Grosset-Fournier & Demachy SARL, 20, rue de Maubeuge, F-75009 Paris (FR).

(81) Designated States *(national)*: CA, JP, US.

(84) Designated States *(regional)*: European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published:
— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NEW POLYPEPTIDES AND THEIR USE FOR THE RESCUE OF FUSION DEFECTIVE VIRUS OR RETROVIRUS GLYCOPROTEINS

(57) Abstract: The present invention relates to the use of at least one activating polypeptide comprising an amino acid sequence corresponding for instance to: the amino acid sequence of the surface subunit (SU) of an envelope glycoprotein of a first virus, or of a first retrovirus; for the rescue of the fusion property of a fusion defective envelope glycoprotein, which has the property of attachment to target cells and is inactive with respect to the fusion with said target cells, said fusion defective envelope glycoprotein belonging to a second virus or a second retrovirus of the same type as or of a different type from the above designated "first virus or first retrovirus".

WO 00/71578 A2

NEW POLYPEPTIDES AND THEIR USE FOR THE RESCUE OF FUSION DEFECTIVE VIRUS OR RETROVIRUS GLYCOPROTEINS

The invention relates to new polypeptides, and their use for the rescue of fusion defective virus or retrovirus glycoproteins.

Retroviruses enter cells following their attachment to specific cell surface receptors. This function is mediated by the envelope glycoproteins expressed as trimeric complexes on the viral particles (Hunter, 1997). Interaction with the receptor is thought to cause a conformational change of the glycoprotein structure necessary to expose a "fusion peptide", buried within the native envelope complex and involved in the actual membrane fusion process (Gerlier, 1998). On the basis of both genetic (Battini et al., 1995; Battini et al., 1992; Morgan et al., 1993; Ott and Rein, 1992) and structural (Fass et al., 1997; Fass et al., 1996) evidence, the attachment and fusion functions of murine leukemia virus (MLV) envelopes have been separated on the two subunits of the envelope monomer. Thus it is generally considered that the amino-terminal half of the MLV SU (surface) subunit is responsible for the binding to the receptor whereas the ectodomain of the TM (transmembrane) subunit contains the "fusion machinery" composed of an amino-terminal fusion peptide associated to a coiled-coil structure that rearrange upon receptor interaction.

How retrovirus envelope glycoproteins convert binding to the receptor to a signal which activates the fusion machinery is currently unknown. Ample evidence from the literature indicate that functional inter-relations between different domains of the envelope complex are essential to promote fusion activation (Lavillette et al., 1998; Rein et al., 1998; Zhao et al., 1997; Zhao et al., 1998).

Retroviruses can be involved in two different types of membrane fusion :

- a fusion between viral and cell membranes triggered by the interaction of the virion envelope glycoproteins with specific cell surface molecules, which are the viral receptors (virus-cell fusion), with said fusion being responsible for retroviral infection,
- a fusion between two cell membranes triggered by the interaction of the envelope glycoproteins expressed at the surface of a cell and the viral receptors expressed on a neighbouring cell (cell-cell fusion), the visible outcome of this latter being the formation of cell syncytia containing up to 100 nuclei (also known as polykaryocytes or multinucleated

giant cells); syncytium-formation results in the death of the cells which make up the syncytium.

Furthermore, cleavage of the R peptide (inhibitory peptide on the cytoplasmic tail of the TM subunit of the envelope glycoprotein) is required to activate fully the fusogenic potential of these envelope glycoproteins and mutants lacking the R peptide show greatly enhanced activity in cell fusion assays (Rein et al, J. Virol., 1994, 68, 1773-1781; Ragheb & Anderson, J. Virol., 1994, 68, 3220-3231; Brody et al, J. Virol., 1994, 68, 4620-4627).

Furthermore, it is known that a vector directing the expression on a eukaryotic cell surface of a syncytium-inducing polypeptide can be used in gene therapy of malignant diseases, as a result of selective elimination of unwanted cells, through syncytia formation (WO 98/40492).

The fusogenicity of envelope glycoproteins has been enhanced as described in Lavillette et al. (1998) and internal application WO 99/36561. The proline-rich region (PRR) of amphotropic MLV envelope glycoprotein was shown to be an essential determinant of membrane fusion activation and specific PRR-mutated envelope glycoproteins were obtained and were characterised by an increased cell-cell or virus-cell fusogenicity.

Studies on Mo-MLV envelope protein GP70 have shown that the first eight amino acids of the mature gp70 protein are essential for fusion and virus entry. Within this region, a particularly important residue is H8. This residue belongs to the PHQ motif which is conserved among all MLVs and some other retroviruses, such as a gibbon ape leukemia virus SEATO and feline leukemia virus types A, B, and C. The amphotropic counterpart of Del2-8, which has the first six amino acids (Met-Ala-Glu-Ser-Pro-His) deleted has been constructed. Vectors bearing the deletion mutant envelope protein could not transduce NIH 3T3 or HeLa cells (Bae et al., Journal of Virology, March 1997, p. 2092-2099).

As the fusion mechanisms induced by interaction of retrovirus envelope glycoproteins and receptors of cell membrane are partially unknown, the regulation of the fusogenicity of retroviral or viral envelope glycoproteins was up to day not possible.

One of the aims of the invention is to enable fusion control of envelope glycoproteins with target cells.

It is an aim of the invention to provide means for the rescue of the fusion property of fusion defective viral or retroviral envelope glycoproteins.

It is another aim of the invention to provide a process for restoring the fusogenic property of retroviral or viral envelope glycoproteins, the fusogenic property of which has been inactivated and for controlling their activation by external means.

It is another aim of the invention to provide a process in which envelope glycoproteins are in a first stage inactivated as to their fusion property and then in a second stage activated in a controlled and regulated way, by means of activating polypeptides, which possibly present targeting capacities, and which are able to restore efficiently the fusogenic and cytotoxic properties of said envelope glycoproteins.

It is another aim of the invention to provide polypeptides liable to activate the cell-cell fusogenicity or the virus-cell fusogenicity.

It is another aim of the invention to provide polypeptides liable to activate the virus-cell fusogenicity and which thus can be used in strategies involving retroviral vectors comprising envelope glycoproteins, the fusion property of which can be activated.

It is another aim of the invention to provide fusogenic envelope glycoproteins presenting cytotoxic properties for use in the treatment of malignant diseases, for which the fusogenic property could be controlled by means of activating polypeptides.

It is another aim of the invention to provide cytotoxic fusogenic envelope glycoproteins having cellular specificity and for which the fusogenic property can be controlled by means of activating polypeptides.

All these aspects are achieved through the present invention, which in one of its broad embodiments is the use of at least one activating polypeptide comprising an amino acid sequence corresponding to :

- the amino acid sequence of the surface subunit (SU) of an envelope glycoprotein of a first virus, or of a first retrovirus, or
- an amino acid sequence derived from said surface subunit, or
- a fragment of an amino acid sequence of said surface subunit, encompassing the receptor binding domain (RBD) and the amino terminal part of the surface subunit (SU),

for the rescue of the fusion property of a fusion defective envelope glycoprotein, which has the property of attachment to target cells and is inactive with respect to the fusion with said target cells, said fusion defective envelope glycoprotein belonging to a second virus or a second retrovirus of the same type as or of a different type from the above designated "first virus or first retrovirus".

Thus the invention lies in the possibility of desactivating the inactivation of the fusogenic property of an envelope glycoprotein, said envelope glycoprotein having previously been made inactive with respect to fusion. This restoration of the fusion property of the envelope glycoprotein is unexpectedly achieved by polypeptides whose common structure is the presence of the receptor binding domain (RBD) and the amino terminal part of SU in the environment of cells expressing fusion defective envelope glycoproteins and of cells expressing retroviral or viral receptors and the interaction of which will lead to the formation of syncytia, or in the environment of retroviruses incorporating fusion defective envelope glycoproteins and of cells expressing retroviral or viral receptors and the interaction of which will lead to infection.

The expression "polypeptide comprising an amino acid sequence corresponding to", means that the polypeptide has the same amino sequence as that of the surface subunit, or as that of an aminoacid sequence derived from the surface subunit or of a fragment and can be prepared by recombinant technique or by chemical synthesis.

The expression "polypeptide comprising an amino acid sequence derived from said surface subunit" means that the polypeptide has an aminoacid sequence which is not strictly identical to that of the sequence of the surface subunits and may contain mutations such as deletion or substitution of aminoacids of the parental sequence or insertion of aminoacid sequence of other origins.

The "amino terminal part of the surface subunit" is defined as the aminoacid sequence of the first aminoacids of a length of about 10 to about 15 aminoacids of the envelope surface subunit, not including the aminoacid sequence of the signal peptide.

A fusion defective envelope glycoprotein derives from a parental envelope glycoprotein in which the fusogenic property has been inactivated, but which still presents the property of recognizing (attachment) target cells.

The expression "inactive with respect to the fusion" means that in a test for the measure of the fusogenic activity, the fusion index is either nul or significantly reduced when the fusion assay is allowed to proceed for about 20 minutes.

The rescue of the fusion property can be measured according to one of the following tests :

- test of formation of syncytia : expressing vectors encoding the envelope glycoprotein to be tested are transfected into cells on day 0. On day 2, the transfected cells are overlaid by indicator cells which express the cell surface receptors corresponding to the tested envelope glycoprotein. Detection of syncytia (i.e. fusion of envelope-expressing cells with neighbouring indicator cells) is performed at day 3 or day 4 and fusion indexes are measured as previously described (Lavillette et al, 1998) by counting the number of syncytia and the number of nuclei outside and inside the syncytia.

- test of infection of target cells : retroviral vectors carrying a lac Z transgene are being generated with the tested envelope glycoproteins. Serial dilutions of the vectors are applied on target cells at day 0. Infected cells are stained at day 2 with X-Gal to reveal the number of cell colonies having integrated the lac Z-vectors and titers are determined according to standard procedures (Lavillette at al., 1998).

It is to be noted that the activating polypeptide corresponds to the amino acid sequence of the SU of an envelope glycoprotein of a first virus or retrovirus and that the fusion defective envelope glycoprotein belongs to a second virus or retrovirus of the same type or of a type different from the first virus or retrovirus. This means that there are in fact two possibilities :

- either the first and second viruses are of the same type, which means that the activating polypeptide and the fusion defective envelope glycoprotein are homologous,
- or the first and second viruses are of a different type, which means that the activating polypeptide and the fusion defective envelope glycoprotein are heterologous.

According to an advantageous embodiment the fusion defective envelope glycoprotein involved are those derived from retroviruses.

Retroviral envelope glycoproteins allow receptor binding of the viral particles onto which they have been incorporated during virion assembly. This preliminary attachment on the viral receptor activates a second function of the retroviral envelope glycoproteins which induces the fusion between the viral and the target cell membranes. While for type C mammalian retroviruses the determinants of receptor binding (RBD, receptor binding domain), are located at the amino-terminal end of the SU envelope subunit, the molecular determinants that are responsible of membrane fusion are located in the TM subunit of the glycoprotein. The molecular determinants or events which are involved in the conversion of binding to the receptor into a signal which activates the TM fusion determinants are not precisely known although by mutagenesis analysis, they are believed to be present in all the SU subunit, in particular in the N-terminal end of the RBD itself (Bae et al., 1997;

Lavillette et al., 2000), in the proline-rich region (Lavillette et al., 1998) and in the carboxy-terminal domain of the SU (Nussbaum et al., 1993; Pinter et al., 1997). Thus the control of the activation of the fusion machinery harbored by the TM subunit might be achieved by introducing specific mutations that affect the fusion-activation determinant of the SU. This invention describes several mutations in the SU which can inhibit or partially inhibit the SU determinants which control activation of the TM fusion-machinery. It is important to understand that such mutations do not impair the fusogenic potential of the mutant envelope glycoproteins, but only impair the molecular determinants that control the activation of the fusion machinery. The latter one is still intact, despite the mutations introduced in the SU, but is not activated upon receptor binding. In addition the invention describes the use of soluble polypeptides that can alleviate the effect of these mutations and hence couple the fusion machinery to a device that is provided exogenously in substitution to the normal fusion control determinants which are disrupted in the mutant envelope glycoproteins. Thus the addition of the "activating" polypeptides can restore or rescue the fusogenicity of the mutant envelope glycoproteins.

Several mutations or modifications of the envelope glycoprotein can be used to reversibly inhibit or reduce their fusogenicity.

A first method is the introduction of the delH mutation that removes an histidine from the PHQ conserved motif located at the amino-terminal end of the SU envelope subunit of type C mammalian retroviruses such as MLV-A, MoMLV and GALV (mutations H5del, H8del or H9del, respectively). Such a mutation induces a drastic effect in the capacity of the receptor binding domain (RBD) of these envelope glycoproteins to triggers further events required to promote membrane fusion.

A second method consists in the replacement of the SU C-terminal domain (C domain) with an heterologous C domain derived from the envelope glycoprotein of another type C mammalian retrovirus. For example; the chimeric CMO amphotropic glycoprotein (hereafter defined) can be constructed by combining the amphotropic MLV RBD and PRR (proline-rich region) with the ecotropic MoMLV C-terminal domain (example 4, Fig. 29A and 29B). For such chimeras, there is a low compatibility between their RBD (amphotropic origin) and C domain (ecotropic origin), which significantly impairs the kinetic of membrane fusion.

A third method consists in the replacement of the RBD by a ligand such as a cytokine or a single chain antibody targeted to a cell surface specific receptor. In contrast to the RBD of the SU of a type C mammalian retrovirus, such ligands do not contain the determinants that allow further activation of the Env fusogenicity. Such chimeric envelope glycoproteins therefore resembles a delH envelope glycoprotein which would carry an RBD that is targeted to a specific receptor.

This invention relates to a method which can alleviate the inhibitory effect of these different mutations and thus restore the fusogenic properties of the mutant envelope glycoproteins. This method uses preparations of polypeptides which are derived from the RBD of different type C mammalian retrovirus envelope glycoproteins. For example, the addition of an RBD which is compatible with the C domain of a fusion-defective envelope glycoprotein will fully restore its fusogenic properties and will activate the kinetic of membrane fusion. Similarly such a peptide will also restore the fusogenicity of envelope glycoproteins that harbor a delH mutation or, alternatively, which carry a cell-type specific ligand in place of the RBD.

This method can be applied to the activation of infectivity of retroviral vectors whose envelope glycoproteins carry one of these fusion-defective mutations. Depending on the presence of a ligand in the fusion-deficient envelope glycoprotein or, alternatively, in the activating polypeptide, control of infection to specific target cell types will therefore be obtained. The invention can also be applied to the control of the hyper-fusogenic activity of retroviral envelope glycoproteins that can induce cell-cell fusion and elimination of unwanted cells by syncytia formation. Indeed while most retroviral envelope glycoproteins are not truly fusogenic when expressed at the surface of the cells that produce the retroviral particles, they acquire their fusion competency during or after assembly and release of the viral particles. Yet there is considerable interest to convert the retroviral envelope glycoproteins into hyperfusogenic glycoproteins that can fuse cells in the close vicinity of the cells into which they are expressed. In addition to the introduction of the above-mentioned mutations that can control activation of their fusogenicity, the fusogenicity of retroviral envelope glycoproteins can be stimulated by several ways to obtain the best effect upon activation with an exogenous activating polypeptide. Thus, additional mutations will be introduced into the retroviral envelope glycoprotein, such as the truncation of the cytoplasmic tail (Rless mutation) and/or the modification of their proline-rich regions.

According to an advantageous embodiment, the activating polypeptide used in the invention recognizes substantially the same receptors of the target cells as the fusion defective envelope glycoprotein.

According to another advantageous embodiment, the activating polypeptide used in the invention recognizes receptors of the target cells which are different from the ones recognized by the fusion defective envelope glycoprotein.

It is thus quite unexpected that the activating polypeptides can restore the fusogenicity of fusion defective envelope glycoproteins by interacting with retroviral receptors which are expressed at the surface of the same target cells, and which are different from the receptors recognized by the fusion defective envelope glycoprotein.

According to another advantageous embodiment of the invention, the activating polypeptide and/or the fusion defective envelope glycoprotein is (are) linked to a ligand, which is specific with respect to some receptors of the target cells, the ligand being preferably linked to the activating polypeptide.

The addition of specific ligands on the activating polypeptides and/or on the fusion defective envelope glycoprotein enables to control the fusion activity of said fusion defective envelope glycoprotein and to increase the specificity with respect to target cells. Therefore, the design of a specific system to eliminate unwanted cells, may consist in a fusion defective envelope glycoprotein associated to an activating polypeptide that is linked to a ligand that can specifically recognize a receptor expressed on the target cells, whereas the activating polypeptide recognizes the same receptor or a different receptor expressed on the target cell.

The link between the activating polypeptide and the ligand is preferably a covalent bond, whereby both the activating polypeptide and the ligand have been genetically fused and encoded by an expression vector as a single protein chain, and whereby the ligand is fused either to the amino-terminus or to the carboxy terminus of the activating polypeptide.

According to another advantageous embodiment, the ligand is chosen among molecules liable to recognize receptors of target cells, for instance receptors of tumoral cells, such as EGF, FGF, VEGF, or liable to recognize epitope of proteins expressed by the target cells, such as recombinant monoclonal antibody.

According to another advantageous embodiment, the activating polypeptide is under the form of a product as such, or is expressed by a host, such as a virus or a cell, or is expressed, in particular by the target cells, transformed by nucleotide sequences coding for said activating polypeptide.

In the use of the invention, one resorts either directly to chemical substances (which is the meaning of activating polypeptide under the form of a product as such), or to host (virus or cell) expressing said activating polypeptide.

In the following of the text, the expression "soluble activating polypeptides" will be sometimes used, meaning that said polypeptides are not integrated in the envelope proteins of a virus.

The capacity of the activating polypeptides to rescue the infectivity of fusion defective envelope glycoproteins is similar whether the target cells i) constitutively express and secrete the activating polypeptides, ii) are co-incubated with both fusion defective envelope glycoprotein and said activating polypeptides iii) are pre-incubated with said activating polypeptides before adding the fusion defective envelope glycoprotein, or iv) are pre-incubated with the fusion defective envelope glycoprotein before adding the activating polypeptides.

In an advantageous use of the invention the fusion defective envelope glycoprotein is expressed by a host, such as a virus incorporating in its envelope said fusion defective envelope glycoprotein, or such as a cell, transformed by nucleotide sequences coding for said fusion defective envelope glycoprotein.

According to the invention, the activating polypeptide advantageously corresponds to a fragment of about 200 to about 250 amino acids of said surface subunit comprising the receptor binding domain (RBD) and comprising the first amino acids of the amino terminal part of the surface subunit, the number of said first aminoacids varying from about 10 to about 15, in particular comprising the RBD and the PHQ peptide, in particular the fragments corresponding to amino acid sequences derived from type C mammalian retroviruses such as:

- Murine leukemia virus such as :
    . amphotropic
    . ecotropic
    . xenotropic
    . polytropic
    . 10 A1 MLV

- or GALV type retroviruses and related retrovirus such as SSAV, FELV-A, FELV-B, FELV-C

in particular wherein the activating polypeptide corresponds to one of the following amino acid sequences :

A-RBD
E-RBD
GALV-RBD
34-1-A-RBD
34-1-E-RBD
34-1-GALV-RBD
EGF-A-RBD
EGF-E-RBD
EGF-GALV-RBD

- A-RBD corresponds to the activating polypeptide comprising the RBD domain and the PHQV motif of the amphotropic MLV surface subunit.

- E-RBD corresponds to the activating polypeptide comprising the RBD domain and the PHQV motif of the ecotropic MLV surface subunit.

- GALV-RBD corresponds to the activating polypeptide comprising the RBD domain and the PHQP motif of the GALV surface subunit.

- 34.1-A-RBD corresponds to the activating polypeptide A-RBD above defined, linked to a ligand which is a single chain antibody directed against MHCI.

- 34.1-E-RBD corresponds to the activating polypeptide E-RBD above defined, linked to a ligand which is a single chain antibody directed against MHCI.

- 34.1 GALV-RBD corresponds to the activating polypeptide GALV-RBD above defined, linked to a ligand which is a single chain antibody directed against MHCI.

- EGF-A-RBD corresponds to the activating polypeptide A-RBD above defined, linked to a ligand which is the epidermal growth factor.

- EFG-E-RBD corresponds to the activating polypeptide E-RBD above defined, linked to a ligand which is the epidermal growth factor.

- EGF-GALV-RBD corresponds to the activating polypeptide GALV-RBD above defined, linked to a ligand which is the epidermal growth factor.

According to the invention, the fusion defective envelope glycoprotein advantageously corresponds to the amino acid sequence of a mutated form of an envelope glycoprotein of a virus or a retrovirus which, in its non mutated form, presents attachment and fusion properties with respect to target cells, said fusion defective envelope glycoprotein comprising one or several mutations in the amino terminal part of its surface subunit, particularly in the first amino acids of the amino terminal part of the surface subunit, the number of said first aminoacids varying from about 10 to about 15, said mutations inactivating the fusion property without substantially affecting the attachment property, in particular comprises one or several mutations within the PHQ amino sequence, wherein in particular the fusion defective envelope glycoprotein is derived from type C mammalian retroviruses such as:

- Murine leukemia virus such as :
  - . amphotropic
  - . ecotropic
  - . xenotropic
  - . polytropic
  - . 10 A1 MLV
- or GALV type retroviruses and related retrovirus such as SSAV, FELV-A, FELV-B, FELV-C,

in particular wherein the fusion defective envelope glycoprotein corresponds to one of the following amino acid sequences :

- MO del H
- AdelH
- PROMOdelH
- GALVdelH
- C2delH
- MO del H R less
- AdelHRless
- PROMOdelHRless
- GALVdelHRless
- C2delHRless
- C MO

- TM MO
- C TM MO
- C MOdelH
- TM MOdelH
- C TM MOdelh
- 34.1-A

* MO del H corresponds to the ecotropic MLV envelope glycoprotein, in which the histidine residue of the PHQV motif is deleted (i.e. the 8th residue of the SU subunit is removed).

* AdelH corresponds to the amphotropic MLV envelope glycoprotein, in which the histidine residue of the PHQV motif is deleted (i.e. the 5th residue of the SU subunit is removed).

* PROMOdelH corresponds to the amphotropic MLV envelope glycoprotein, in which the proline rich region has been replaced by the proline rich region of the ecotropic MLV envelope and in which the histidine residue of the PHQV motif is deleted.

* GALVdelH corresponds to the GALV envelope glycoprotein, (gibbon ape leukemia virus envelope glycoprotein), in which the histidine residue of the PHQP motif is deleted (i.e. the 9th residue of the SU subunit is removed).

* C2delH corresponds to the amphotropic MLV envelope glycoprotein, in which the proline rich region has been modified by a small deletion at its carboxy terminus and in which the histidine residue of the PHQV motif is deleted (i.e. the 5th residue of the SU subunit is removed).

* MO del H R less corresponds to the ecotropic MLV envelope glycoprotein, in which the histidine residue of the PHQV motif is deleted and in which the R peptide of the TM subunit has been deleted.

* AdelHRless corresponds to the amphotropic MLV envelope glycoprotein in which the histidine residue of the PHQV motif is deleted and in which the R peptide of the TM subunit has been deleted.

* PROMOdelHRless corresponds to the amphotropic MLV envelope glycoprotein, in which the proline rich region has been replaced by the proline rich region of the ecotropic MLV envelope, in which the histidine residue of the PHQV motif is deleted and in which the R peptide of the TM subunit has been deleted.

* GALVdelHRless corresponds to the GALV envelope glycoprotein, in which the histidine residue of the PHQP motif is deleted and in which the R peptide of the TM subunit has been deleted.

* C2delHRless corresponds to the amphotropic MLV envelope glycoprotein, in which the proline rich region has been modified by a small deletion at its carboxy terminus and in which the histidine residue of the PHQV motif is deleted (i.e. the 5th residue of the SU subunit is removed) and in which the R peptide of the TM subunit has been deleted.

* CMO envelope corresponds to the amphotropic MLV envelope glycoprotein, in which the SU carboxy-terminal domain, C domain, has been replaced by the C domain derived from the ecotropic MoMLV envelope glycoprotein.

* CMOdelH envelope corresponds to the amphotropic MLV envelope glycoprotein, in which the SU carboxy-terminal domain, C domain, has been replaced by the C domain derived from the ecotropic MoMLV envelope glycoprotein and in which the histidine residue of the PHQV motif is deleted (i.e. the 5th residue of the SU subunit is removed).

* TMMO envelope corresponds to the amphotropic MLV envelope glycoprotein, in which the ectodomain of the TM subunit has been replaced by the TM ectodomain derived from the ecotropic MoMLV envelope glycoprotein.

* TMMOdelH envelope corresponds to the amphotropic MLV envelope glycoprotein, in which the ectodomain of the TM subunit has been replaced by the TM ectodomain derived from the ecotropic MoMLV envelope glycoprotein and in which the histidine residue of the PHQV motif is deleted (i.e. the 5th residue of the SU subunit is removed).

* CTMMO envelope corresponds to the amphotropic MLV envelope glycoprotein, in which both the SU carboxy-terminal domain, C domain, and the ectodomain of the TM subunit have been replaced by the C domain and TM ectodomain derived from the ecotropic MoMLV envelope glycoprotein.

* CTMMOdelH envelope corresponds to the amphotropic MLV envelope glycoprotein, in which both the SU carboxy-terminal domain, C domain, and the ectodomain of the TM subunit have been replaced by the C domain and TM ectodomain derived from the ecotropic MoMLV envelope glycoprotein and in which the histidine residue of the PHQV motif is deleted (i.e. the 5th residue of the SU subunit is removed).

* 34.1-A envelope corresponds to the amphotropic MLV envelope glycoprotein, in which the receptor binding domain has been removed and replaced by an single chain antibody against human class I MHC molecules.

The invention also relates to the use of at least one activating polypeptide and of at least one fusion defective envelope glycoprotein for the preparation of a drug useful for the treatment of cancer and infectious diseases.

The invention also relates to products containing an activating polypeptide and a fusion defective envelope glycoprotein as defined above, as a combined preparation for simultaneous, separate, or sequential use for the rescue of the fusion defective property of said fusion defective envelope glycoprotein.

The invention also relates to products containing an activating polypeptide and a fusion defective envelope glycoprotein as defined above, as a combined preparation for simultaneous, separate or sequential use for the treatment of cancer pathologies.

The invention also relates, as new products, to activating polypeptides corresponding to a fragment of about 200 to about 250 amino acids of said surface subunit comprising the RBD and the PHQV peptide, in particular the fragments corresponding to amino acid sequences derived from type C mammalian retroviruses such as:

- Murine leukemia virus such as :

. amphotropic
. ecotropic
. xenotropic
. polytropic
. 10 A1 MLV

- or GALV type retroviruses and related retrovirus such as SSAV, FELV-B

in particular wherein the activating polypeptide corresponds to one of the following amino acid sequences :

A-RBD
E-RBD
GALV-RBD
34-1-A-RBD
34-1-E-RBD

34-1-GALV-RBD
EGF-A-RBD
EGF-E-RBD
EGF-GALV-RBD

The invention also relates, as new products, to fusion defective envelope glycoprotein to the amino sequence of an envelope glycoprotein of a virus or a retrovirus which in its native state presents attachment and fusion properties with respect to target cells, and comprises one or several mutations in the amino terminal part of its surface subunit, inactivating the fusion property without substantially affecting the attachment property, in particular comprises one or several mutations within the PHQ amino sequence of the amino terminal part of its surface subunit, wherein in particular the fusion defective envelope glycoprotein is derived from type C mammalian retroviruses such as:

- Murine leukemia virus such as :
  . amphotropic
  . ecotropic
  . xenotropic
  . polytropic
  . 10 A1 MLV
- or GALV type retroviruses and related retrovirus such as SSAV, FELV-A, FELV-B or FELV-C

in particular wherein the fusion defective envelope glycoprotein corresponds to one of the following amino acid sequences :

- AdelH
- PROMOdelH
- GALVdelH
- C2delH
- AdelHRless
- PROMOdelHRless
- GALVdelHRless
- C2delHRless
- C MO
- TM MO

- C TM MO
- C MOdelH
- TM MOdelH
- C TM MOdelh
- 34.1-A

with the proviso that they are different from the ecotropic MLV sequence, wherein the amino acids in position 2 to 8 are either substituted or deleted.

The invention also relates to the use of at least one activating polypeptide to the rescue of the fusion property of a fusion defective envelope glycoprotein wherein the fusion defective envelope glycoprotein corresponds to the amino acid sequence of a mutated form of an envelope glycoprotein of a virus or a retrovirus which, in its non mutated form, presents attachment and fusion properties with respect to target cells, said fusion defective envelope glycoprotein being advantagously derived from type C mammalian retroviruses such as:

- Murine leukemia virus such as :
  . amphotropic
  . ecotropic
  . xenotropic
  . polytropic
  . 10 A1 MLV
- or GALV type retroviruses and related retrovirus such as SSAV, FELV-A, FELV-B, FELV-C and comprising :

* one or several mutations in the amino terminal part of its surface subunit, particulary in the first amino acids of the amino terminal part of the surface subunit, the number of said first aminoacids varying from about 10 to about 15, said mutations inactivating the fusion property without substantially affecting the attachment property, in particular comprises one or several mutations within the PHQ amino-terminal sequence, wherein in particular the fusion defective envelope glycoprotein corresponds to one of the following amino acid sequences :

- MO R less del H
- MO del H

- A R less del H
- AdelH
- PROMO R less del H
- PROMOdelH
- GALV R less del H
- GALVdelH
- C2 R less del H
- C2delH
- C MOdelH
- TM MOdelH
- C TM MOdelh

* or the replacement of the amino-terminal domain of its surface subunit (RBD) by a ligand such as a growth factor, or a cytokine or a single chain antibody, the extent of said mutations inactivating the fusion property and the attachement property to the normal viral receptor, in particular wherein the fusion defective envelope glycoprotein corresponds to the following amino-acid sequence :

- 34.1-A

* or the replacement of its SU carboxy-terminal domain and / or its TM ectodomain, and preferably its SU carboxy-terminal domain, by the corresponding SU carboxy-terminal domain and / or the TM ectodomain of another "second" glycoprotein derived from type C mammalian retroviruses, with the proviso that said fusion defective glycoprotein and said second glycoprotein are not derived from the same type C mammalian retroviruses, in particular wherein the fusion defective envelope glycoprotein corresponds to one of the following amino-acid sequences :

- CMO
- TM MO
- C TM MO

FIGURE LEGENDS

Fig 1A, 1B and 1C represent titres of AdelH retroviruses in the presence of soluble envelope fragments of the invention. Retroviral vectors carrying VSV-G or/and wild-type amphotropic, A, envelope glycoproteins were used as controls. Titres expressed as lacZ infectious units (i.u.) per ml of viral supernatants. In figures 1A, 1B and 1C, the titres are on the y-axis.

Fig 1A. Different target cell types (XC, NIH3T3, CHO-PIT-2, TE 671, from left to right) were incubated with both lacZ retroviruses and conditioned media containing, or not (control), the indicated polypeptides during the 3 hrs of infection. The black columns correspond to the control, the grey columns correspond to A-RBD and the hatched columns correspond to SU polypeptide.

Fig 1B. XC target cells were incubated with dilutions of A-RBD (x axis) and infected with LacZ retroviruses carrying the AdelH envelope glycoproteins (curve with black circles), A envelope glycoproteins (curve with black squares) or VSV-G envelope glycoproteins (curve with hollow squares).

Fig 1C. Retroviral vectors carrying A envelope glycoproteins (curves with black squares) or AdelH envelope glycoproteins (curves with black circles) were incubated with XC target cells for 1 hr at 4°C to allow virion binding while preventing cell entry. After PBS washing to remove unbound retroviruses, cells were incubated at 37°C with undiluted A-RBD-containing supernatants for the indicated times of dilution represented on the x axis, the titres being expressed on the y axis. A-RBD was then eliminated from the cell supernatant by washing the cells four times with 1 ml PBS (resulting in dilution of unbound A-RBD by more than 100,000 times). Cells were then grown in regular medium for 2 days before X-Gal staining.

Fig 2. represents the titers of A-RBD-mediated activation of AdelH retroviruses. A-RBD (used undiluted) was bound to XC target cells at 37°C for 30 min. Unbound A-RBD was removed and cells were washed two times with PBS and incubated at 37°C. At the indicated times, cells were further washed two times with PBS and incubated at 37°C for 3 hrs with retroviruses carrying A envelope glycoproteins or AdelH envelope glycoproteins. Titers, determined 2 days later by X-Gal staining, are expressed as lacZ i.u./ml (y-axis), as a function of time before infection (x-axis). The curve with black circles corresponds to retrovirus carrying A envelope glycoproteins and the curve with black squares corresponds to retrovirus carrying AdelH envelope glycoproteins.

Fig 3. represents titers of AdelH retroviruses after ultrafiltration. Titers expressed on the y-axis (as lacZ i.u./ml) of retroviruses carrying wild-type, A, or AdelH envelope glycoproteins before (1) or after incubation with A-RBD (2), and following A-RBD elution on two successive 700 KD-cut off ultrafiltration cartridges (3). After elution AdelH retroviruses were subjected to re-stimulation with fresh A-RBD (4).

Fig. 4A and 4B represent infection assays of AdelH and MOdelH retroviruses in the presence of RBD or RBDdelH polypeptides.

Fig. 4A. Retroviruses carrying A (upper left) or AdelH (upper right) envelope glycoproteins were mixed with A-RBD (hatched column) or A-RBDdelH (grey column) during infection of XC cells. Retroviruses carrying MO (lower left) or MOdelH (lower right) envelope glycoproteins were mixed with E-RBD (hatched column) or E-RBDdelH (grey column) during infection of XC cells. Titers are expressed as lacZ i.u./ml on the y-axis. In Figure 4A, the black column corresponds to control.

Fig. 4B. represents binding assays of A-RBD (solid lines) and A-RBDdelH fragments (broken lines) on Cear13 (top), and binding assays of E-RBD (solid lines) and E-RBDdelH fragments (broken lines) on Cear13 cells (bottom). The background of fluorescence (filled histograms) was provided by incubating the cells with supernatants devoid of envelope fragments.

Fig 5A and 5B represent the cross-activation of AdelH and MOdelH retroviruses by MLV RBDs.

Fig. 5A. Titers of retroviruses carrying AdelH or MOdelH envelope glycoproteins in the presence of the indicated RBDs. Titers are determined (y-axis) on XC target cells as lacZ i.u./ml. On figure 5A, the black column corresponds to control, the hatched column corresponds to A-RBD and the grey column corresponds to E-RBD.

Fig. 5B. Titers of AdelH retroviruses (y-axis) on CHO cells expressing, or not (control), PiT-2 and/or mCAT-1 MLV receptors in the presence of the indicated RBDs. On figure 5B,

the hatched column corresponds to E-RBD, the grey column corresponds to A-RBD and the black column corresponds to control.

Fig. 6 represents the A-RBD sequence

Fig. 7 represents the E-RBD sequence

Fig. 8 represents the GALV-RBD sequence

Fig. 9 represents the EGF-A-RBD sequence

Fig. 10 represents the EGF-E-RBD sequence

Fig. 11 represents the EGF-GALV-RBD sequence

Fig. 12 represents the 34.1-A-RBD sequence

Fig. 13 represents the 34.1-E-RBD sequence

Fig. 14 represents the 34.1-GALV-RBD sequence

Fig. 15 represents the AdelH sequence

Fig. 16 represents the PROMOdelH sequence

Fig. 17 represents the GALVdelH sequence

Fig. 18 represents the C2delH sequence

Fig. 19 represents the AdelHRless sequence

Fig. 20 represents the PROMOdelHRless sequence

Fig. 21 represents the GALVdelHRless sequence

Fig. 22 represents the C2delHRless sequence

Fig. 23 represents the C MO sequence

Fig. 24 represents the C MO delH sequence

Fig. 25 represents the TM MO sequence

Fig. 26 represents the TM MOdelH sequence

Fig. 27 represents the C TM MO sequence

Fig. 28 represents the C TM MO delH sequence

Fig. 29A represents the schematic structure of chimeric amphotropic envelope glycoproteins in which the C domain (CMO chimera), the TM ectodomain (TMMO chimera) or both C and TM domains (CTMMO chimera) are replaced by the ecotropic counterpart derived from MoMLV envelope glycoprotein. The CMOdelH, TMMOdelH and CTMMOdelH chimeric envelope glycoproteins harbor the H5del mutation.

Figure 29B represents the kinetic of infection of retroviral vectors carrying the wild-type amphotropic envelope glycoprotein (A) or the CTMMO amphotropic chimera on XC cells in the presence, or not, of soluble polypeptides encompassing the amphotropic RBD (A-RBD) or the ecotropic RBD (E-RBD), as indicated.

The curve with points and black circles corresponds CTMMO on XC cells + E-RBD

The curve with points and hollow circles corresponds to A on XC cells + E-RBD

The curve with continuous line and hollow circles corresponds to A on XC cells

The curve with dotted line and hollow circles corresponds to A on XC cells + A-RBD

The curve with continuous line and black circles corresponds to CTMMO on XC cells

The curve with dotted line and black circles corresponds to CTMMO on XC cells + A-RBD

The X axis corresponds to the time expressed in seconds.

The Y axis corresponds to the number of infected cells.

Fig. 30 represents the 34.1-A Sequence

Example 1:

MATERIALS AND METHODS

Cell lines. The TELCeB6 cell line (Cosset et al., 1995b) has been derived from the TELac2 line after transfection and clonal selection of a Moloney murine leukemia virus (MoMLV)-based expression plasmid to produce Gag and Pol proteins. The TELac2 cells were originally derived from the TE671 human rhabdomyosarcoma cells (ATCC CRL8805) to express the nlsLacZ reporter retroviral vector (Takeuchi et al., 1994). Production of infectious retroviral particles by TELCeB6 cells depends on newly introduced envelope expression vectors.

Cerd9 and Cear13 cells (Kozak et al., 1995) and CHO-PiT-2 cells (Rodrigues and Heard, 1999) were derived from CHO (Chinese hamster ovary) cells (ATCC CCL-61) and express ecotropic MLV receptors alone or both ecotropic and amphotropic receptors or amphotropic MLV receptors alone, respectively.

XC-A-RBD cells were derived from XC rat sarcoma cells (ATCC CCL-165) after transfection with the pA-ST plasmid expressing the amino terminal receptor binding domain of the amphotropic envelope glycoprotein (Battini et al., 1996).

TE671 and TELCeB6 cells were grown in DMEM (Life-Technologies) supplemented with 10% fetal bovine serum. NIH-3T3 mouse fibroblasts were grown in DMEM supplemented with 10% new-born calf serum. Cerd9, Cear13, CHO-PiT-2 and CHO cells were grown in DMEM supplemented with 10% fetal bovine serum and with proline (Life-Technologies).

Construction of envelope expression vectors. Plasmids FBASALF and FBMOSALF carrying a phleomycin-resistance gene and encoding the MLV-4070A amphotropic (noted as A) and MoMLV ecotropic (noted as MO) envelope glycoproteins, respectively, have been described elsewhere (Cosset et al., 1995a) and were used as backbones for construction and expression of envelope mutants. All constructs were generated by PCR-mediated and oligonucleotide site-directed mutagenesis (details and sequences available upon request).

The FBASALF plasmid was modified to produce a cell entry-defective form of the amphotropic glycoprotein, designated AdelH envelope, by deleting the 36th codon of the 4070A env gene (Ott et al., 1990). The resulting mutant envelope glycoprotein, in which the 5th residue of the SU envelope subunit was removed, was named AdelH. The expression plasmid encoding the fusion-defective MOdelH envelope glycoproteins, harboring a deletion of the 8th residue of the SU envelope subunit corresponding to the 41th codon of the MoMLV env gene (Shinnick et al., 1981), was derived from FBMOSALF.

Plasmids encoding soluble receptor binding domains (RBD) were derived from FBASALF and FBMOSALF expression vectors. The boundaries of either amphotropic (A-RBD) and ecotropic (E-RBD) RBDs, defined as A32244 and A34269, were fused in frame with the carboxy-terminal ends of either envelope glycoproteins by fusion to aminoacids R623 or R634, respectively. Residues are numbered starting from the initiation methionine deduced from the aminoacid sequences of the 4070A amphotropic MLV (Ott et al., 1990) or the Moloney MLV (Shinnick et al., 1981). Expression vectors encoding either A-RBDdelH or E-RBDdelH were generated similarly by using the plasmids expressing the AdelH or MOdelH envelope glycoproteins. In plasmids encoding A-RBD, E-RBD, A-RBDdelH or E-RBDdelH polypeptides, the 11 aminoacid-long VSV tag (YTDIEMNRLGK) (Kreis and Lodish, 1986) was inserted after either RBDs.

Transfections, production of RBDs and infection assays. Envelope glycoproteins expression plasmids were transfected into TELCeB6 as reported elsewhere (Cosset et al., 1995a). RBD expression vectors were transfected by calcium phosphate precipitation in NIH-3T3, XC or TE671 cells. Transfected cells were selected with phleomycin (50 µg/ml) and phleomycin-resistant colonies were pooled.

RBD-containing supernatants were collected after an overnight production from confluent RBD-transfected cells, filtrated trough 0.45 micron pore-sized membranes and stored at 4°C. Virus-containing supernatants were collected after an overnight production from freshly confluent env-transfected TELCeB6 cells in regular medium.

The supernatant were filtered trough 0.45 micron pore-sized membranes and diluted in DMEM for titration assays. Target cells were seeded in 24 multi-well plates at a density of $5x10^4$ cells per well and incubated overnight at 37°C. Unless otherwise indicated in figure legends, 200 µl of conditioned cell culture media containing the RBDs were added to the cells after removing their supernatants. Then 200 µl of viral supernatant dilutions containing 5 µg/ml polybrene were added and cells were incubated for 3 hrs at 37°C. Cell supernatants were then removed and cells were incubated in regular medium for 48 hrs. X-Gal staining and viral titer determination were performed as previously described (Cosset et al., 1995a) and expressed as lacZ i.u. (infectious units)/ml.

Antibodies. Anti-SU: a rat monoclonal antibody 83A25 (Evans et al., 1990) cell culture supernatant against MLV SU used undiluted for FACS analysis. Anti-VSV tag: a purified mouse monoclonal antibody P5D4 (SIGMA-Aldrich) used diluted 1/100 for FACS analysis.

Binding assays and FACS analysis. For binding assays, target cells were washed in PBS and detached by a 10 min incubation at 37°C with versene 0.02% in PBS. Cells were washed in PBA (PBS with 2% FCS and 0.1% sodium azide). $5x10^5$ cells were incubated with 1 ml of conditioned supernatants containing the VSV tagged RBDs for 45 min at 37°C. Cells were then washed with PBA and were incubated with the P5D4 antibody, for 45 min at 4°C. Cells were washed twice with PBA and incubated with anti-mouse FITC-conjugated antibodies (DAKO, U.K.), respectively. 5 min before the two final washes in PBA, cells were counter-stained with 20 µg/ml propidium iodide. Fluorescence of living cells was analysed with a fluorescent-activated cell sorter (FACSCalibur, Beckton Dickinson).

RESULTS

H5del phenotype can be efficiently compensated by soluble A-RBD. Viral particles were generated with either wild-type amphotropic A or AdelH envelope glycoproteins and used to infect target cells in the presence or in the absence of soluble SU or soluble A-RBD, an envelope fragment encompassing the amphotropic receptor binding domain. When SU or A-RBD were provided during infection the infectivity of retroviruses carrying wild-type amphotropic envelope glycoproteins was decreased by approximately 5-100 times (Fig. 1A), most likely owing to the partial blocking of PiT-2 receptors on the target cell surface. In striking contrast, while in the absence of envelope fragments retroviruses generated with AdelH envelopes were very poorly infectious, with titers in the range of 10^1-10^2 lacZ i.u. (infectious units)/ml, the presence of SU or of A-RBD in the medium of infected cells could strongly stimulate the infectivity of the former retroviruses by up to 30,000 fold (Fig. 1A). Stimulation of AdelH infection with either soluble envelope fragments was detected on a wide range of target cell types including rat XC cells, 3T3 murine fibroblasts, TE671 human cells and PiT-2-transfected CHO hamster cells, though with different efficacies (Fig. 1A). Significant differences were detected when soluble SU or A-RBD were used to activate the AdelH retroviruses. Comparatively to soluble SU, A-RBD was 100 to 1,000 fold more potent to activate AdelH retroviruses (Fig. 1A). These results indicated that the determinants of activation were located in the first half of the SU protein. Therefore to facilitate the characterization of this activation pathway, the subsequent experiments were performed by using the soluble RBD. A-RBD could fully activate AdelH retroviruses when diluted 10 fold and retained 90% of activity at a 1:100 dilution, thus suggesting that it was effective at very low doses (Fig. 1B). In contrast the infectivity of control retroviruses pseudotyped with either wild-type amphotropic envelopes or with VSV-G glycoproteins was weakly increased or unchanged, respectively, when A-RBD was diluted (Fig. 1B). Additionally the origin of the cell types used to produce the envelope fragments did not influence the stimulation of infectivity since conditioned media harvested from 3T3, TE671 or XC cells that were engineered to express the envelope fragments could similarly activate the infectivity of AdelH retroviruses (data not shown). Finally activation of AdelH retroviruses was found to be the result of a specific interaction of A-RBD with the PiT-2 amphotropic receptor. Indeed while A-RBD could efficiently rescue the infectivity of AdelH retroviruses on PiT-2-transfected CHO cells, no infection could be detected when using parental CHO cells as target cells (see below and Fig. 5A). Altogether these data demonstrated that activation of AdelH retroviruses by A-RBD was specific.

The capacity of the soluble envelope fragments to rescue the infectivity of AdelH retroviruses was similar whether the target cells i) constitutively expressed and secreted the envelope fragments, ii) were co-incubated with both AdelH retroviruses and A-RBD, iii) were pre-incubated with A-RBD prior to infection, or iv) were pre-incubated with the AdelH retroviruses before adding A-RBD (Table 1). Using this latter experimental condition, the kinetic of activation of infection was found very rapid. Indeed a brief incubation of target cells with A-RBD, by no more than 30 sec, was found sufficient to fully stimulate the infectivity of pre-bound AdelH retroviruses (Fig. 1C).

A-RBD stimulates the target cell membrane rather than the AdelH envelope. In theory, it is possible that the envelope fragments might rescue the infectivity of AdelH retroviruses either by interacting with the fusion-defective viral particles or by activating the target cell membrane upon receptor binding. Indeed within the envelope complex, the SU subunits are not tightly attached to either the TM subunits or to the other SU units and can easily shed off the viral particles (Hunter and Swanstrom, 1990). However the reassociation of shed SU with the viral particle is unlikely since they are incorporated on virions by virtue of their association with the TM subunit that results from their synthesis as a common SU-TM polypeptide precursor (Hunter and Swanstrom, 1990). Therefore it is unlikely that soluble SU or A-RBD mixed with AdelH retroviruses may interact with the envelope complex of the latter retroviruses and thus rescue their fusion defect. Nevertheless the following experiments were performed to address this possibility.

A-RBD was bound to target cells at 37°C for 30 min. After removal of unbound fragments by washing target cells, binding of A-RBD was verified by FACS analysis. Cells were then incubated at 37°C for 0 (T0) to 28 hrs (T28) to allow internalisation of receptor/A-RBD complexes (Rodrigues and Heard, 1999) and A-RBD disappearance from the cell surface, as checked by FACS analysis (Data not shown). Cells were then washed again to remove A-RBD that may have been released by the cells and the AdelH retroviruses were then added at the different time points. Comparatively to infection performed at T0, no significant decrease of infectivity could be detected when the AdelH retroviruses were added until 8 hrs after the initial binding with A-RBD. Likewise the infectivity of AdelH retroviruses added at T20 was at least 50 fold higher than before pre-stimulation of the cells with A-

RBD, thus indicating that stimulation of infectivity was durable (Fig. 2). Since A-RBD was efficiently internalised and removed from the cell surface after 4 hrs of incubation (Data not shown), these results indicated that activation of AdelH retroviruses most likely occurred by activation of the cell surface rather than by interaction with the AdelH envelope complex itself.

Further evidence indicated that A-RBD could not associate with the AdelH retroviruses. Indeed retroviruses generated with either A or AdelH envelopes and mixed or co-expressed with A-RBD were not found to incorporate A-RBD as shown by western-blotting of purified viral particles with antibodies recognising A-RBD (data not shown). However these experiments could not exclude that undetected quantities of A-RBD might be associated with the virions. Therefore to rule out this possibility, a mixture of A-RBD and retroviruses carrying either AdelH or wild-type A envelope glycoproteins were separated on two consecutive 700 KD-pore sized ultrafiltration columns in order to elute out A-RBD. No decrease of infectivity could be detected for retroviruses carrying wild-type amphotropic envelopes after the two subsequent ultrafiltrations (Fig. 3), demonstrating that this process did not affect the viability of the retroviruses. However while before ultrafiltration the AdelH viral particles mixed with A-RBD could efficiently infect XC target cells, as expected, only a residual infectivity could be detected after A-RBD elution (Fig. 3), indicating that stimulation of AdelH retroviruses was lost after separation from A-RBD. As control, the AdelH retroviruses processed through the two columns retained their full capacity to be stimulated by newly added A-RBD (Fig. 3). Similar experiments were conducted after separation of AdelH retroviruses from A-RBD by sephacryl chromatography and led to the same conclusions (data not shown).

Altogether these data indicated that stimulation of infectivity of the AdelH retroviruses proceeded via activation of the target cell membrane, most likely as a consequence of interaction of A-RBD with the PiT-2 amphotropic receptor, rather than via interaction with the AdelH envelope incorporated on the retrovirus itself. Thus A-RBD could activate the target cell membrane which in turn became competent to allow entry of retroviruses carrying the fusion-defective AdelH envelope glycoproteins.

The SU amino-terminal end activates a cell entry pathway common to type C mammalian retroviruses. The histidine residue deleted in the H5del mutant glycoprotein belongs to a peptide motif, PHQ, found at the amino-termini of the SUs of all type C mammalian retroviruses, suggesting a conserved function of this motif. Thus the H8del mutant ecotropic envelope glycoprotein (MOdelH), harboring the deletion of the histidine in the PHQV peptide motif of the MoMLV SU, is impaired for both cell-cell and virus-cell fusion (Bae et al., 1997). Similarly the H9del mutation of the GALV envelope glycoprotein GAdelH mutant results in impairement of both cell-cell and virus-cell fusion. Further experiments were performed to demonstrate that MOdelH, AdelH and GAdelH (delH) envelopes were phenotypically similar. Indeed the E-RBD fragment, containing the MoMLV ecotropic receptor binding domain, could rescue the infectivity of MOdelH retroviruses by up to 1,000 fold (Fig. 4A). Similar results were obtained for GAdelH retrovirus, whose fusogenicity could be efficiently rescued by soluble envelope fragment encompassing the GALV envelope glycoprotein receptor-binding domain. This result indicated that fusion activation by soluble envelope fragments was not particular to AdelH envelopes and suggested that the MLV SU amino-terminus may convey a signal common to other type C mammalian retrovirus which may be required to achieve early post-binding events.

To address this hypothesis the delH mutations, H5del, H8del, were introduced in the A-RBD or in the E-RBD fragments, respectively. The resultant A-RBDdelH and E-RBDdelH could bind their respective cell surface receptors as efficiently as the parental envelope fragments (Fig. 4B). They could also decrease the infectivity of retroviruses bearing wild-type amphotropic or ecotropic glycoproteins, respectively, by up to 10-fold (Fig. 4A). This effect was most likely a consequence of partial receptor blocking. However in contrast to A-RBD or to E-RBD, neither the A-RBDdelH nor the E-RBDdelH fragments could rescue the infectivity of AdelH or MOdelH retroviruses, respectively (Fig. 4A). These data therefore demonstrated that the integrity of the amino-terminal ends of either RBD fragments was absolutely required to activate postbinding functions and suggested that an essential determinant of cell membrane activation may reside at the amino-terminal end of the MLV RBD. Thus the so-called "receptor binding domain" located in the amino-terminal half of the MLV SU may in fact be composed of two different entities, one involved in specific receptor binding and a second, likely to be non functional in AdelH or MOdelH envelopes, involved in transmission of a signal that may activate cell entry upon receptor binding.

Therefore to test whether the binding function and the activating function of the MLV receptor binding domain could be uncoupled, cells "infected" with AdelH or MOdelH retroviruses were cross-incubated with either E-RBD or A-RBD polypeptides, respectively. Interestingly the infectivity of MOdelH retroviruses could be activated by either A-RBD or E-RBD fragments (Fig. 5A). This cross-activation was reciprocal since either fragments could activate, with similar efficacies, the infectivity of AdelH retroviruses (Fig. 5A). The cross-reactivity of either RBDs was dependent on the presence of both mCAT-1 and PiT-2 receptors on the cell surface since neither PiT-2-transfected CHO cells nor mCAT-1-transfected CHO cells could be infected when E-RBD was used to activate the infectivity of AdelH retroviruses (Fig. 5B) and vice-versa (data not shown). The specificity of activation of AdelH/MOdelH retroviruses by ecotropic envelope fragments was further demonstrated by using the MOD84K-ST envelope fragment, harboring the D84K point mutation that inactivates MoMLV envelope binding (MacKrell et al., 1996). This fragment could activate neither AdelH nor MOdelH retroviruses despite the presence of both amphotropic and ecotropic receptors on target cells (Fig. 5A). Cross-activation of AdelH retroviruses could also be obtained by using soluble envelope glycoproteins from Gibbon Ape leukemia virus (data not shown), thus indicating that infection by delH retroviruses could be activated by a pathway common to type C mammalian retrovirus.

Table 1. Activation of AdelH retroviruses by A-RBD

Cells	Pre-binding[a]	binding and infection[b]	Titre[c]
XC	-	lacZ(AdelH)	$5x10^1$
XC	-	lacZ(AdelH)[d] + A-RBD[d]	$3.1x10^5$
XC-A-RBD[e]	-	lacZ(AdelH)	$9x10^4$
XC	A-RBD	lacZ(AdelH)	$1.1x10^5$
XC	lacZ(AdelH)	A-RBD	$2.7x10^5$

a: Target cells were pre-incubated for one hr at 4°C with conditioned supernatants containing A-RBD (used undiluted) or with lacZ retroviral vectors carrying AdelH envelope glycoproteins, as indicated.

b: Target cells were incubated for 3 hrs at 37°C.

c: Titres as lacZ i.u./ml determined 48 hrs post-infection.

d: Both A-RBD (used undiluted) and AdelH retroviruses were added at the same time.

e: XC cells constitutively expressing A-RBD polypeptides.

Example 2: Control of syncytia formation by RBD polypeptides

Ideally, the transfer of therapeutic genes into tumour cells should lead both to direct local cell destruction and to activation of anti-tumour immunity to clear tumour deposits to which the genes cannot be delivered. The most commonly used genes for the control of local tumour growth have been the suicide gene/prodrug systems such as Herpes Simplex Virus thymidine kinase (HSVtk)/Ganciclovir (GCV) or cytosine deaminase (CD)/5FC systems. Importantly, these systems not only kill gene-transduced target cells but also close neighbours via local bystander effects. These effects depend upon cell-cell communication or on the diffusion of toxic metabolites, but generally act only over short ranges. Nonetheless, the existence of such bystander effects are crucial to compensate for the relatively poor efficiencies of gene transfer that are currently possible. However, results from early clinical trials have demonstrated that both better delivery systems and more effective genes are clearly still required to improve the efficiency of both local tumour control and of anti tumour immune responses. Hence, therapeutic genes with greater local killing activity would represent a significant advance for gene transfer strategies for cancer.

It has been recently demonstrated that the well-known ability of several viral fusogenic membrane glycoproteins (FMG) to fuse target cells could be exploited therapeutically to kill tumour cells. These previous results introduced the FMG as a novel class of therapeutic genes which are more potent than existing suicide gene systems in killing tumour cells and do so by mechanisms which are highly immunostimulatory. Table 2 shows the results of four FMGs derived from the non fusogenic ecotropic or amphotropic murine leukemia virus glycoproteins, whose fusogenicity were increased by deleting the p2R peptide (naturally removed during retrovirus egress) for the ARless and MoRless FMG and also by modifying the proline-riche region (Lavillette et al., 1998) for the C2Rless and PROMORless FMGs.

The four resulting FMGs, Arless, C2Rless and PROMORless and MORless induced syncytia formation in most of the cell types that express amphotropic receptors and/or ecotropic receptors. However for selective eliminations of cells as required in cancer gene therapy protocols, the design of a mechanism that can control the formation of syncyia is required. We therefore introduced in all four FMGs the H5del or the H8del mutation (delH mutation) which removes the 5th or the 8th residue of the amphotropic or the ecotropic SU, respectively, an histidine belonging to the PHQV conserved motif. This mutation could inactivate an early event of the fusion process after receptor binding (Example 1 and Lavillette et al., 2000). As previously demonstrated the virus-cell fusion activity of delH MLV envelopes bearing the delH mutation can be efficiently rescued by using polypeptides encompassing the first 213 amino-acids of amphotropic MLV SU (A-RBD) or the first 236 amino-acids of ecotropic MoMLV SU (E-RBD) (example 1 and Lavillette et al., 2000). Consistently with Example 1, the data in Table 2 demonstrated that the cell-cell fusion activity of the four FMGs engineered to harbor the delH mutation, AdelRless, C2delHRless, and PROMOdelHRless and MOdelHRless, could be rescued by providing either the A-RBD or E-RBD polypeptides thus leading to syncytia formation and elimination of target cells that express the relevant viral receptors. Interestingly formation of syncytia by delH FMGs could be achieved by using an heterologous polypeptide (E-RBD) that interacted with a receptor different of that recognised by the delH FMGs.

Table 2. Modulation of cell-cell fusion by soluble envelope fragments

Envelope	RBD	Syncytia	delH envelope	RBD	Syncytia
A	-	-			
ARless	-	++	AdelHRless	-	-
				A-RBD	+/-
				E-RBD	+
C2Rless	-	+++	C2delHRless	-	-
				A-RBD	+/-
				E-RBD	++
PROMORless	-	++++	PROMOdelHRless	-	+/-
				A-RBD	++
				E-RBD	++++
MORless	-	++++	MOdelHRless	-	-
				A-RBD	++++
				E-RBD	++++

fusion indexes (calculated according to Lavillette et al., 1998) -: less than 100, +/-: 100-304, +: 300-1000, ++: 1000-10 000, +++: 10 000-15 000, ++++: >15000

Of note, similar results of fusion activation with RBD polypeptides were obtained with the GALVdelHRless envelope glycoprotein, indicating that the delH mutation of the potent GALVRless FMG (Bateman et al., 2000) can be controlled by adding a soluble Env-derived polypeptide (Lavillette and Cosset, unpublished results). These data indicate that it is possible to exploit a well-known phenomenon - the fusion of virally infected cells with uninfected cells as a novel, potent cytotoxic therapy using gene transfer. There are several options to exploit the use of delH FMGs in cell elimination stategies. Firstly viral vectors such as adenovirus vectors can be used to transduce both the delH FMG and the gene encoding the activating RBD polypeptide in cancer cells in vivo. Secondly cells transduced in vitro or ex vivo by a vector expressing a delH fusion-defective FMG can be induced to form syncytia after inoculation of an RBD presented as a protein preparation. In addition as well as using direct gene transfer, delH FMG and RBD can be effectively delivered through cell-mediated mechanisms, which will also be useful as an option for *in vivo* delivery of FMG into tumours for both local tumour eradication as well as enhancement of anti tumour vaccination strategies.

Example 3: Addition of ligands on RBD polypeptides

To further control the fusion activation of FMGs by RBD polypeptides, a ligand was inserted at the amino-terminal end of the RBD. The anti-MHC-I 34.1 scFv was chosen as we previously demonstrated that it could redirect the host-range of retroviral vectors carrying 34.1-ecotropic MLV chimeric envelope glycoproteins and allow specific infection of human cells (Marin et al. 1996). Therefore the 34.1 scFV was fused to A-RBD or to E-RBD and preparation of 34.1-A-RBD or 34.1-E-RBD polypeptides were obtained. Retroviral vectors carrying wild-type envelope glycoproteins, A, or AdelH fusion-defective amphotropic envelope glycoproteins were used to infect target cells in the presence of A-RBD, E-RBD, 34.1-E-RBD or 34.1-A-RBD polypeptides (Table 3). While either polypeptides could only weakly decrease the infectivity of parental retroviruses, they could strongly stimulate the infectivity of AdelH retrovirus. No significant differences could be detected when comparing the A-RBD, the E-RBD, the 34.1-E-RBD or the 34.1-A-RBD polypeptides in activation of AdelH retrovirus infectivity (Table 3).

Table 3. Modulation of virus-cell fusion by ligand-displaying envelope fragments

Envelope	RBD fragment	Titers
A	-	8.1×10^6
	A-RBD	5.6×10^5
	E-RBD	8×10^6
	34.1-A-RBD	1.35×10^6
	34.1-E-RBD	3×10^6
AdelH	-	2×10^1
	A-RBD	8.5×10^4
	E-RBD	8×10^5
	34.1-A-RBD	7.3×10^4
	34.1-E-RBD	7×10^5

These data therefore suggest that it is possible to control the fusion activity of FMGs by adding specific ligands on the activating polypeptides. Therefore the design of a specific system to eliminate unwanted cells may consist in a delH FMG associated to an RBD polypeptide that displays a ligand that can specifically recognise a receptor expressed on the target cells. Table 4 shows that 34.1-A-RBD and 34.1-E-RBD could activate the cell-cell fusion activity of the PROMOdelHRless FMG with an efficiency similar to that of A-RBD or E-RBD.

Altogether these results indicate that since the 34.1-E-RBD or 34.1-A-RBD polypeptides could cross-activate the fusogenicity of AdelH, AdelHRless, or PROMOdelHRless envelope glycoproteins on cells that express the targeted antigen (the MHC class I molecules) but no receptor corresponding to either the chimeric RBD or the delH retrovirus (Tables 3 & 4), efficient retargeted infection or cell-type specific syncytia formation could be obtained.

Table 4. Modulation of cell-cell fusion by ligand-displaying envelope fragments

Envelope	RBD	Syncytia
A	-	18
ARless	-	4464
ARlessdelH	-	14
	RBD	1014
PROMORless	-	14796
PROMORlessdelH	-	304
	A-RBD	18920
	E-RBD	17000
	34.1-A-RBD	19000
	34.1-E-RBD	20000

Example 4: Activation of Env fusogenicity by RBD and C/TM domains interaction

The H5del mutation was introduced into the CMO, the TMO and the CTMMO MLV amphotropic envelope glycoprotein (Lavillette et al., in preparation), derived from the MLV 4070A envelope glycoprotein by substitution of the SU carboxy-terminal domain (C domain) and/or TM ectodomain with those derived from the MLV ecotropic envelope glycoprotein (Figure 29A). Recent evidence from the literature suggest that the C domain of MLV envelope glycoproteins is involved in activation and/or control of the fusion machinery contained in the TM subunit of the envelope glycoprotein (Nussbaum et al., 1993; Pinter et al., 1997). Retroviruses carrying either the parental or the delH-mutated CMOdelH, TMOdelH and CTMMOdelH chimeric envelope glycoproteins were prepared and their infectivity was assayed in the presence or the absence of A-RBD or E-RBD soluble polypeptides (Table 5). While the infectivity of AdelH control retroviruses could easily be rescued when either A-RBD or E-RBD polypeptides were present in the infection mixture (Lavillette et al., 2000), the infectivity of CMOdelH, TMMOdelH and CTMMOdelH retroviruses could not be rescued with the A-RBD polypeptide although it was efficiently rescued with the E-RBD polypeptide (Table 5). These results indicated that the fusogenicity of the CMOdelH, TMMOdelH and CTMMOdelH envelope glycoproteins could only be activated by soluble RBD polypeptides derived from the same envelope glycoprotein from which the C and/or TM domain were inserted in the delH chimera.

Further evidence for the requirement of a compatibility between the RBD and the C and/or TM domains were provided by the finding that the CTMMO retroviruses displayed a significantly slower kinetic of infection comparatively to that of retroviruses carrying wild-type MLV amphotropic envelope glycoproteins (Figure 29B). In the absence of exogenous RBD in the infection mixture, the replication retroviruses carrying the CTMMO envelope glycoproteins was delayed, most likely owing to the unability of the MLV 4070A-derived RBD of the latter to activate their MoMLV-derived C and/or TM domains. Yet the addition in the infection mixture of exogenous RBD polypeptides derived from the ecotropic MoMLV envelope glycoprotein (E-RBD), matching the C and/or TM domains of the CTMMO chimera could restore the infection kinetic of CTMMO retroviruses to a level similar if not identical to that of wild-type amphotropic retroviruses (Figure 29B). In contrast the kinetic of infection of retroviruses carrying wild-type amphotropic envelope glycoproteins was not increased in the presence of the E-RBD polypeptide. Additionally, infectivity with retroviruses carrying either wild-type amphotropic or CTMMO envelope glycoproteins was strongly decreased or inhibited, respectively, when infection was performed in the presence of A-RBD (Figure 29B). This inhibitory effect was most likely due to A-RBD-mediated partial blocking of amphotropic receptors in the case of wild-type amphotropic envelope glycoproteins and both receptor interference and incompatibility between this MLV amphotropic Env-derived fragment and the ecotropic C domain in the case of CTMMO chimera.

Table 5

Infectivity in the presence of :

no RBD	A-RBD	E-RBD
10^6	10^4	10^6
10^1	10^5	10^5
10^6	10^2	10^6
10^0	10^0	10^5
10^6	10^3	10^6
10^0	10^3	10^5
10^6	10^1	10^6
10^0	10^0	10^5
10^6	10^6	10^6

Example 5: Retargeted gene delivery using RBD fragments.

We have therefore demonstrated that the RBD of a retroviral envelope glycoprotein is not only a domain responsible of attaching the viral particle to a cell surface receptor but is also a domain that is responsible, upon receptor binding, of fusion activation most likely via an interaction with the carboxy-terminal domain of the SU Env subunit (C domain). The RBD determinant responsible of this activation appears to be located at its amino-terminal end and can be disrupted by the delH mutation. Thus receptor binding of RBD most probably promotes a conformational change of this amino-terminal end which, in turn, is rendered competent for inducing further conformational changes in the Env C domain that control the fusion machinery located in the TM Env subunit. We surmised that the replacement of the RBD of a retroviral envelope glycoprotein by a cell type specific ligand may redirect the attachment of viral particles carrying such targeting envelope glycoproteins, yet specific attachment of these virions would not promote fusion activation due to the absence in the novel ligand of a molecular device responsible of fusion activation. Such an activation, leading to infection, may therefore be restored by providing in trans a soluble RBD in the infection mixture.

Data in Table 6 demonstrate that retroviral vectors carrying a chimeric envelope glycoprotein in which the wild-type RBD is replaced by a single-chain antibody against human class I MHC molecules cannot infect human cells although they can efficiently bind to class I MHC molecules expressed at the surface of the target cells. The addition in trans of RBD polypeptides during infection can promote the events necessary to allow infection of the target cells, demonstrating that retargeted binding of chimeric viral particles can give rise to productive infection and/or gene delivery when the fusion domains of the targeting envelope are activated by an exogenous polypeptide.

Table 6. RBD-mediated activation of the fusion activity of retargeted envelope glycoproteins

Enva	-	+ RBDb	+ RBD & competitor[c]
A	106	105	105
34.1-A	<101	105	102

a: LacZ retroviral vectors were generated with the indicated envelope glycoproteins. A, wild-type amphotropic Env; 34.1-A, a chimeric envelope glycoprotein derived from the MLV amphotropic Env in which the receptor binding domain was replaced by a single chain antibody (scFv) against human class I MHC molecule (Marin et al, 1995). The vectors were used to infect class I positive human cells. Titers as lacZ i.u./ml.

b: A preparation of A-RBD polypeptide was added during infection. Similar results were obtained when the A-RBD polypeptide was added before or after infection with the 34.1-A retroviruses.

c: Anti-MHC-I IgGs, corresponding to the Mab which was used to derive the anti-MHC-I scFv inserted into the 34.1-A chimera, were used as competitor in this infection assay. The anti-MHC-I IgGs were used to block the targeted antigens before infection with the 34.1-A retroviruses.

REFERENCES

Bae, Y., Kingsman, S.M. and Kingsman, A.J. (1997) *J. Virol.*, 71, 2092-2099.

Bateman, A., F. Bullough, S. Murphy, L. Emiliusen, D. Lavillette, F.-L. Cosset, R. Cattaneo, S. J. Russell, and R. Vile. 2000. Fusogenic Membrane Glycoproteins as a Novel Class of Genes for the Local and Immune-mediated Control of Tumour Growth. Cancer Research.

Battini, J.L., Danos, O. and Heard, J.M. (1995) *J. Virol.*, 69, 713-719.

Battini, J.L., Heard, J.M. and Danos, O. (1992) *J Virol*, 66, 1468-1475.

Battini, J.L., Rodrigues, P., Müller, R., Danos, O. and Heard, J.-M. (1996) *J. Virol.*, 70, 4387-4393.

Carr, C.M., Chaundhry, C. and Kim, P.S. (1997) *PNAS*, 23, 14306-14313.

Cosset, F.-L., Morling, F.J., Takeuchi, Y., Weiss, R.A., Collins, M.K.L. and Russell, S.J. (1995a) *J. Virol.*, 69, 6314-6322.

Cosset, F.-L., Takeuchi, Y., Battini, J.L., Weiss, R.A. and Collins, M.K.L. (1995b) *J. Virol.*, 69, 7430-7436.

Evans, L.H., Morrison, R.P., Malik, F.G., Portis, J. and Britt, W.J. (1990) *J. Virol.*, 64, 6176-6183.

Fass, D., Davey, R.A., Hamson, C.A., Kim, P.S., Cunningham, J.M. and Berger, J.M. (1997) *Science*, 277, 1662-1666.

Fass, D., Harrison, S.C. and Kim, P.S. (1996) *Nature Structural Biology*, 3, 465-469.

Gerlier, D. (1998) *Virologie*, 2, 215-226.

Haywood, A. (1994) *J Virol*, 68, 1-5.

Hunter, E. (1997) In J. M. Coffin and S. H. Hughes (eds.), *Viral entry and receptors*. Cold Spring Harbor Laboratory Press, New York, USA, pp. 71-120.

Hunter, E. and Swanstrom, R. (1990) *Curr. Top. Microbiol. Immun.*, 157, 187-253.

Jones, J.S. and Risser, R. (1993) *J. Virol.*, 67, 67-74.

Kizhatil, K. and Albritton, L.M. (1997) *J Virol*, 71, 7145-7156.

Klasse, P.J., Bron, R. and Marsh, M. (1998) *Advanced drug delivery reviews*, 34, 65-91.

Kozak, S.L., Siess, D.C., Kavanaugh, M.P., Miller, A.D. and Kabat, D. (1995) *J. Virol.*, 69, 3433-3440.

Kreis, T.E. and Lodish, H.F. (1986) *Cell*, **46**, 929-937.

Lavillette, D., Maurice, M., Roche, C., Russell, S.J., Sitbon, M. and Cosset, F.-L. (1998) *J Virol*, **72**, 9955-9965.

Lavillette, D., A. Ruggieri, S. J. Russell, and F.-L. Cosset. 2000. Activation of a cell entry pathway common to type C mammalian retroviruses by soluble envelope fragments. Journal of Virology. 74:295-304.

MacKrell, A.J., Soong, N.W., Curtis, C.M. and Anderson, W.F. (1996) *J. Virol*, **70**, 1768-1774.

Marin, M., F. Brockly, D. Noel, M. Etienne-Julan, M. Biard-Piechaczyk, T. D. Hua, Z. J. Gu, and M. Piechaczyk. 1995. Cloning and expression of a single-chain antibody fragment specific for monomorphic determinant of class I molecules of the human major histocompatability complex. Hybridoma. 14:443-451.

Morgan, R.A., Nussbaum, O., Muenchau, D.D., Shu, L., Couture, L. and Anderson, W.F. (1993) *J Virol*, **67**, 4712-4721.

Nussbaum, O., A. Roop, and W. F. Anderson. 1993. Sequences determining the pH dependence of viral entry are distinct from the host range-determining region of the murine ecotropic and amphotropic retrovirus envelope proteins. J. Virol. 67:7402-7405.

Ott, D., Friedrich, R. and Rein, A. (1990) *J Virol*, **64**, 757-766.

Ott, D. and Rein, A. (1992) *J Virol*, **66**, 4632-4638.

Park, B.H., Matuschke, B., Lavi, E. and Gaulton, G.N. (1994) *J Virol*, **68**, 7516-7524.

Pinter, A., Kopelman, R., Li, Z., Kayman, S.C. and Sanders, D.A. (1997) *J Virol*, **71**, 8073-8077.

Ragheb, J.A. and Anderson, W.F. (1994) *J Virol*, **68**, 3220-3231.

Rein, A., Mirro, J., Haynes, J.G., Ernst, S.M. and Nagashima, K. (1994) *J Virol*, **68**, 1773-1781.

Rein, A., Yang, C., Haynes, J.A., Mirro, J. and Compans, R.W. (1998) *J Virol*, **72**, 3432-3435.

Rodrigues, P. and Heard, J.M. (1999) *J. Virol.*, 3789-3799.

Shinnick, T.M., Lerner, R.A. and Sutcliffe, J.G. (1981) *Nature (London)*, **293**, 543-548.

Siess, D.C., Kozak, S.L. and Kabat, D. (1996) *J. Virol.*, **70**, 3432-3439.

Takeuchi, Y., Cosset, F.L., Lachmann, P.J., Okada, H., Weiss, R.A. and Collins, M.K.L. (1994) *J. Virol.*, 68, 8001-8007.

Wang, H., Paul, R., Burgeson, R., Keene, D. and Kabat, D. (1991b) *J Virol*, 65, 6468-6477.

Zhao, Y., Lee, S. and Anderson, W.F. (1997) *J Virol*, 71, 6967-6972.

Zhao, Y., Zhu, L., Benedict, C.A., Chen, D., Anderson, W.F. and Cannon, P.M. (1998) *J Virol*, 72, 5392-5398.

Zhu, N.L., Cannon, P.M., Chen, D. and Anderson, W.F. (1998) *J Virol*, 72, 1632-1639.

CLAIMS

1. Use of at least one activating polypeptide comprising an amino acid sequence corresponding to :

- the amino acid sequence of the surface subunit (SU) of an envelope glycoprotein of a first virus, or of a first retrovirus, or
- an amino acid sequence derived from said surface subunit, or
- a fragment of an amino acid sequence of said surface subunit, encompassing the receptor binding domain (RBD) and the amino terminal part of the surface subunit,

for the rescue of the fusion property of a fusion defective envelope glycoprotein, which has the property of attachment to target cells and is inactive with respect to the fusion with said target cells, said fusion defective envelope glycoprotein belonging to a second virus or a second retrovirus of the same type as or of a different type from the above designated "first virus or first retrovirus".

2. Use of an activating polypeptide according to claim 1, wherein the activating polypeptide recognizes substantially the same receptors of the target cells as the fusion defective envelope glycoprotein or wherein the activating polypeptide recognizes receptors of the target cells which are different from the ones recognized by the fusion defective envelope glycoprotein.

3. Use according to any one of the claims 1 or 2, wherein the activating polypeptide and/or the fusion defective envelope glycoprotein is (are) linked to a ligand, which is specific with respect to some receptors of the target cells, the ligand being preferably linked to the activating polypeptide.

4. Use according to claim 3, wherein the ligand is chosen among molecules liable to recognize receptors of target cells, for instance receptors of tumoral cells, such as EGF, FGF, VEGF.

5. Use according to any one of claims 1 to 4, wherein the activating polypeptide is under the form of a product as such, or is expressed by a host, such as a virus or a cell, or is expressed, in particular by the target cells, transformed by nucleotide sequences coding for said activating polypeptide.

6. Use according to any one of claims 1 to 5, wherein the fusion defective envelope glycoprotein is expressed by a host, such as a virus incorporating in its envelope said fusion defective envelope glycoprotein, or such as a cell, transformed by nucleotide sequences coding for said fusion defective envelope glycoprotein.

7. Use according to any one of claims 1 to 6, wherein the activating polypeptide corresponds to a fragment of about 200 to about 250 amino acids of said surface subunit comprising the receptor binding domain (RBD) and comprising the first amino acids of the amino terminal part of the surface subunit, the number of said first aminoacids varying from about 10 to about 15, in particular comprising the RBD and the PHQ peptide, in particular the fragments corresponding to amino acid sequences derived from type C mammalian retroviruses such as:

- Murine leukemia virus such as :
    . amphotropic
    . ecotropic
    . xenotropic
    . polytropic
    . 10 A1 MLV
- or GALV type retroviruses and related retrovirus such as SSAV, FELV-A, FELV-B, FELV-C

in particular wherein the activating polypeptide corresponds to one of the following amino acid sequences :

A-RBD
E-RBD
GALV-RBD
34-1-A-RBD

34-1-E-RBD
34-1-GALV-RBD
EGF-A-RBD
EGF-E-RBD
EGF-GALV-RBD

8. Use according to any one of claims 1 to 7, wherein the fusion defective envelope glycoprotein corresponds to the amino acid sequence of a mutated form of an envelope glycoprotein of a virus or a retrovirus which, in its non mutated form, presents attachment and fusion properties with respect to target cells, said fusion defective envelope glycoprotein being advantagously derived from type C mammalian retroviruses such as:

- Murine leukemia virus such as :
  . amphotropic
  . ecotropic
  . xenotropic
  . polytropic
  . 10 A1 MLV
- or GALV type retroviruses and related retrovirus such as SSAV, FELV-A, FELV-B, FELV-C and comprising :

* one or several mutations in the amino terminal part of its surface subunit, particulary in the first amino acids of the amino terminal part of the surface subunit, the number of said first aminoacids varying from about 10 to about 15, said mutations inactivating the fusion property without substantially affecting the attachment property, in particular comprises one or several mutations within the PHQ amino-terminal sequence, wherein in particular the fusion defective envelope glycoprotein corresponds to one of the following amino acid sequences :

- MO R less del H
- MO del H
- A R less del H
- AdelH

- PROMO R less del H
- PROMOdelH
- GALV R less del H
- GALVdelH
- C2 R less del H
- C2delH
- C MOdelH
- TM MOdelH
- C TM MOdelh

* or the replacement of the amino-terminal domain of its surface subunit (RBD) by a ligand such as a growth factor, or a cytokine or a single chain antibody, the extent of said mutations inactivating the fusion property and the attachement property to the normal viral receptor, in particular wherein the fusion defective envelope glycoprotein corresponds to the following amino-acid sequence :

- 34.1-A

* or the replacement of its SU carboxy-terminal domain and / or its TM ectodomain, and preferably its SU carboxy-terminal domain by the corresponding SU carboxy-terminal domain and / or the TM ectodomain of another "second" glycoprotein derived from type C mammalian retroviruses, with the proviso that said fusion defective glycoprotein and said second glycoprotein are not derived from the same type C mammalian retroviruses, in particular wherein the fusion defective envelope glycoprotein corresponds to one of the following amino-acid sequences :

- CMO
- TM MO
- C TM MO

9. Use of at least one activating polypeptide and of at least one fusion defective envelope glycoprotein for the preparation of a drug useful for the treatment of cancer and infectious diseases.

10. Products containing an activating polypeptide and a fusion defective envelope glycoprotein as defined in any one of claims 1 to 9, as a combined preparation for simultaneous, separate, or sequential use for the rescue of the fusion defective property of said fusion defective envelope glycoprotein.

11. Products containing an activating polypeptide and a fusion defective envelope glycoprotein as defined in any one of claims 1 to 9, as a combined preparation for simultaneous, separate or sequential use for the treatment of cancer pathologies.

12. Activating polypeptides corresponding to a fragment of about 200 to about 250 amino acids of said surface subunit comprising the RBD and the PHQ peptide, in particular the fragments corresponding to amino acid sequences derived from type C mammalian retroviruses such as:

- Murine leukemia virus such as :

. amphotropic
. ecotropic
. xenotropic
. polytropic
. 10 A1 MLV

- or GALV type retroviruses and related retrovirus such as SSAV, FELV-A, FELV-B, FELV-C

in particular wherein the activating polypeptide corresponds to one of the following amino acid sequences :

A-RBD
E-RBD
GALV-RBD
34-1-A-RBD
34-1-E-RBD
34-1-GALV-RBD
EGF-A-RBD
EGF-E-RBD
EGF-GALV-RBD

13. Fusion defective envelope glycoprotein to the amino sequence of an envelope glycoprotein of a virus or a retrovirus which in its native state presents attachment and fusion properties with respect to target cells, and comprises one or several mutations in the amino terminal part of its surface subunit, inactivating the fusion property without substantially affecting the attachment property, in particular comprises one or several mutations within the PHQ amino sequence of the amino terminal part of its surface subunit, wherein in particular the fusion defective envelope glycoprotein is derived from type C mammalian retroviruses such as:

- Murine leukemia virus such as :
  - . amphotropic
  - . ecotropic
  - . xenotropic
  - . polytropic
  - . 10 A1 MLV
- or GALV type retroviruses and related retrovirus such as SSAV, FELV-A, FELV-B or FELV-C

in particular wherein the fusion defective envelope glycoprotein corresponds to one of the following amino acid sequences :

- AdelH
- PROMOdelH
- GALVdelH
- C2delH
- A R less del H
- PROMO R less del H
- GALV R less del H
- C2 R less del H
- C MO
- TM MO
- C TM MO
- C MOdelH
- TM MOdelH

- C TM MOdelh
- 34.1-A

with the proviso that they are different from the ecotropic MLV sequence, wherein the amino acids in position 2 to 8 are either substituted or deleted.

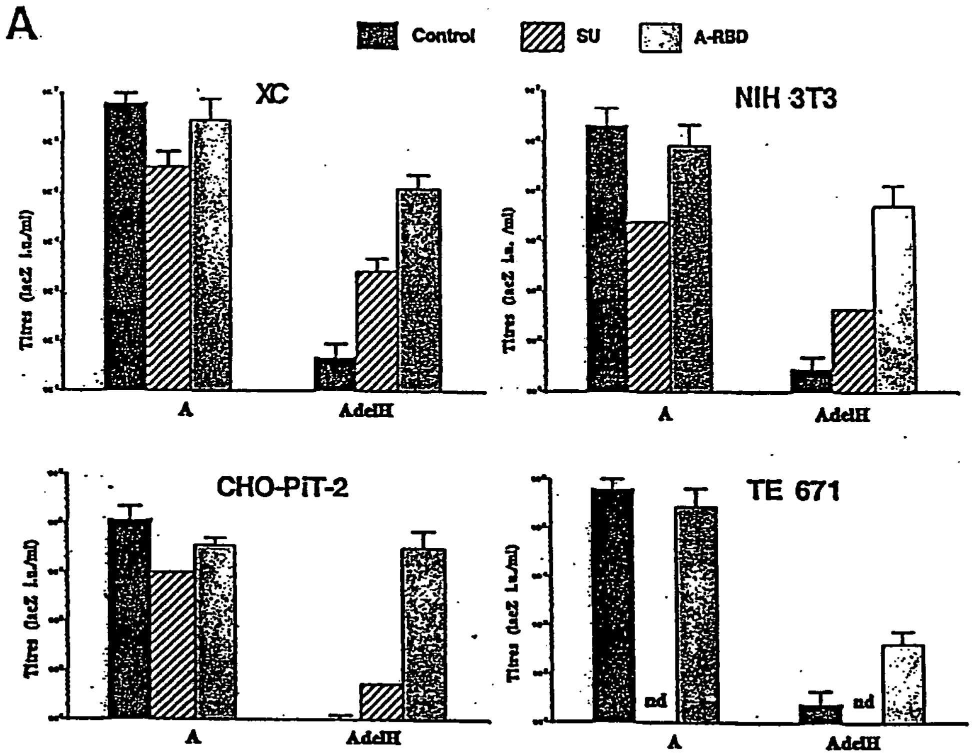

FIG.1A

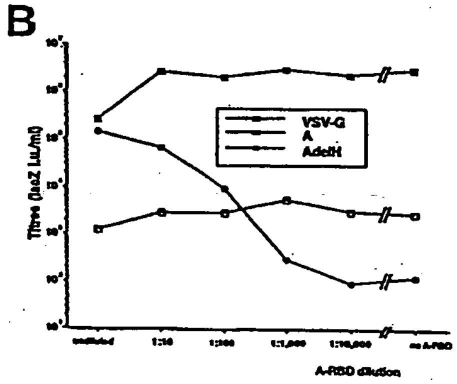

FIG. 1B

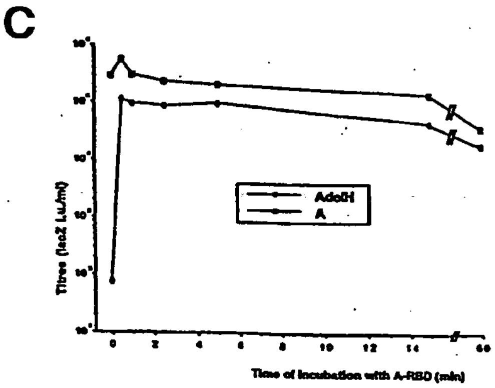

FIG.1C

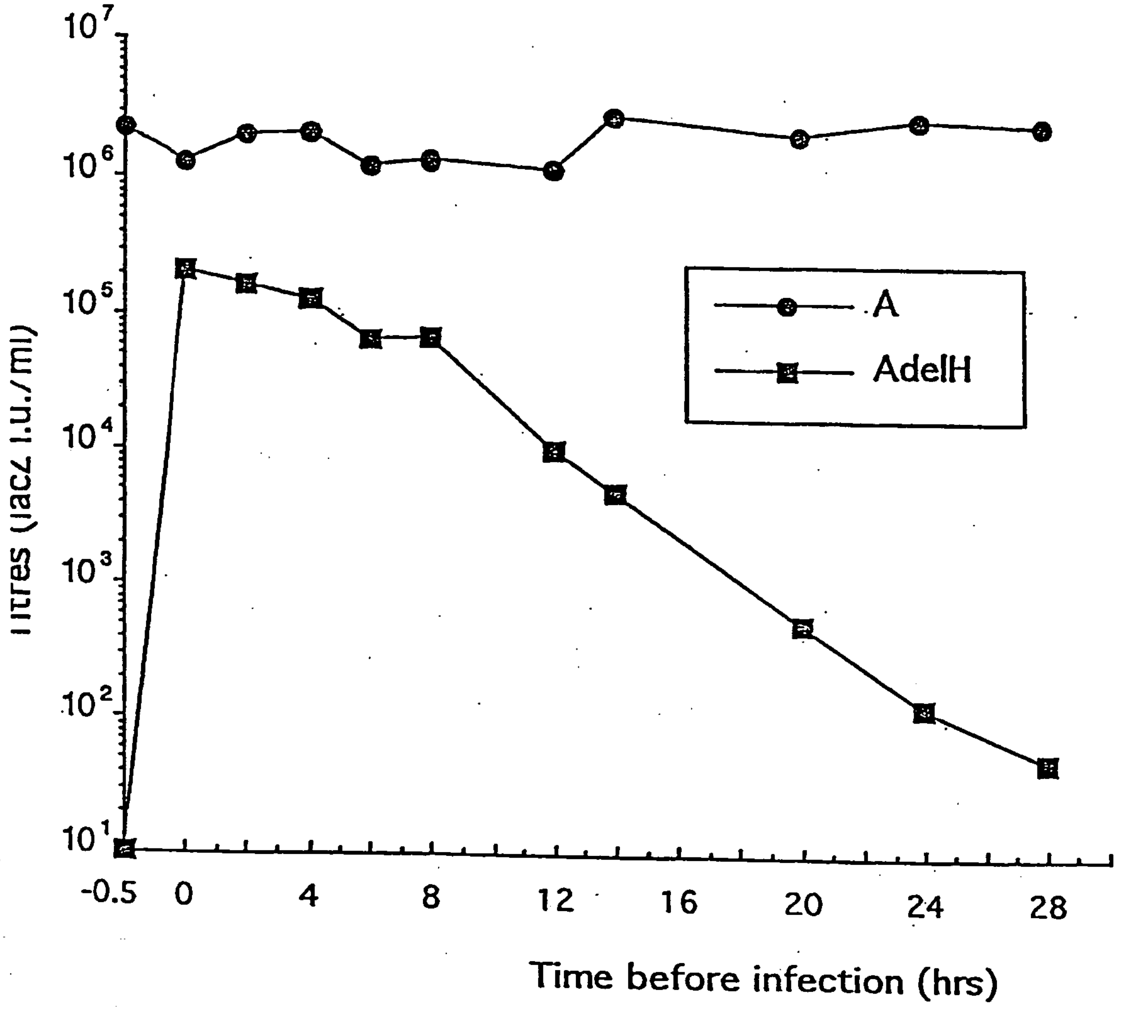
10⁷
10⁶
10⁵
10⁴
10³
10²
10¹
Titres (lacZ i.u./ml)
-0.5
0
4
8
12
16
20
24
28
Time before infection (hrs)
A
AdelH

FIG.2

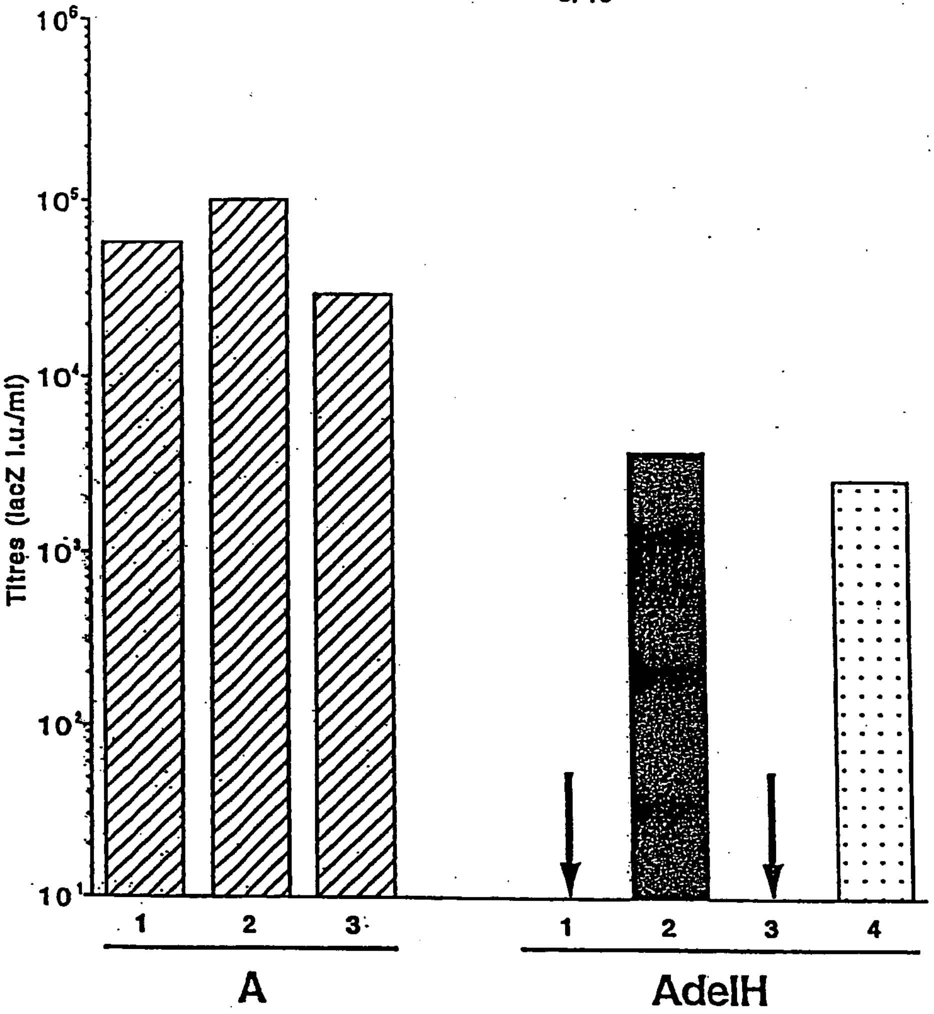

FIG.3

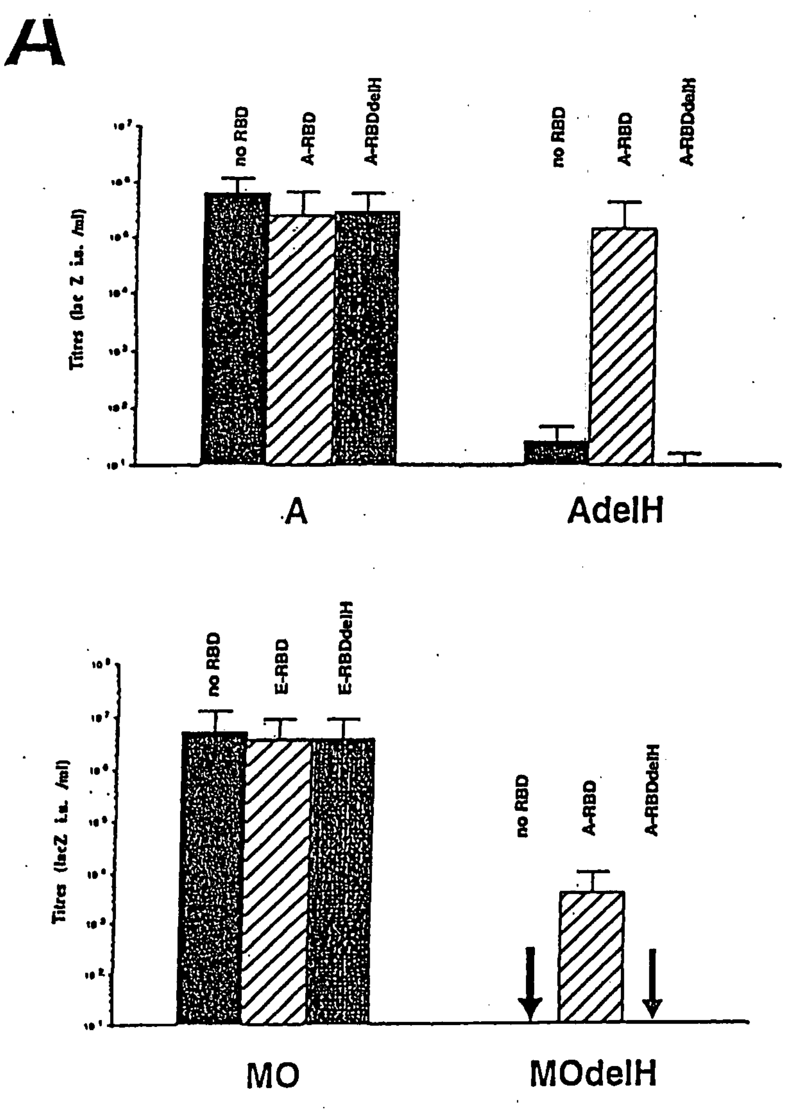
A
Titres (lac Z i.u. /ml)
no RBD
A-RBD
A-RBDdelH
no RBD
A-RBD
A-RBDdelH
A
AdelH
Titres (lacZ i.u. /ml)
no RBD
E-RBD
E-RBDdelH
no RBD
A-RBD
A-RBDdelH
MO
MOdelH

FIG. 4A

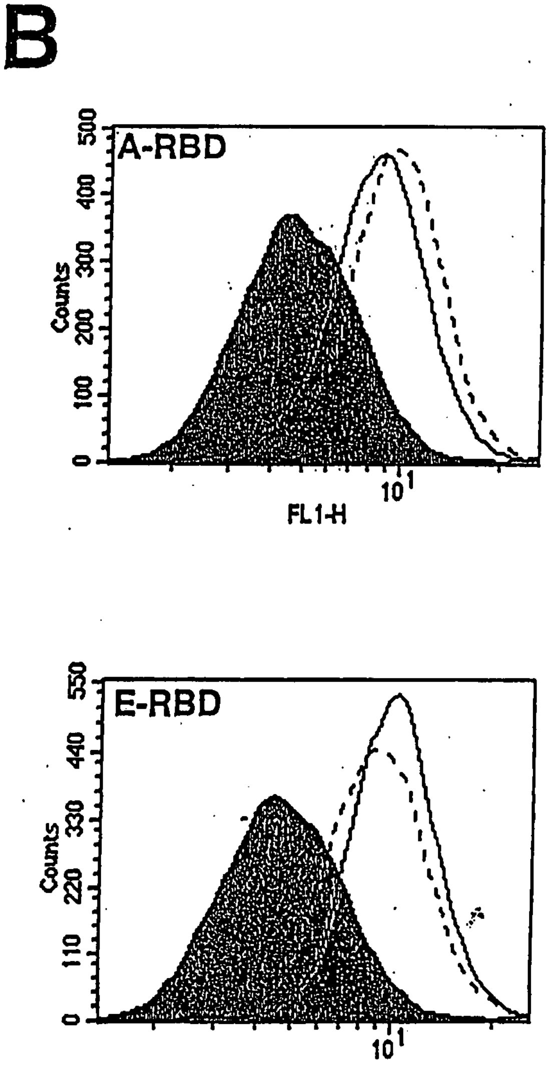
B
A-RBD
Counts
0
100
200
300
400
500
10^1
FL1-H
E-RBD
Counts
0
110
220
330
440
550
10^1

FIG. 4B

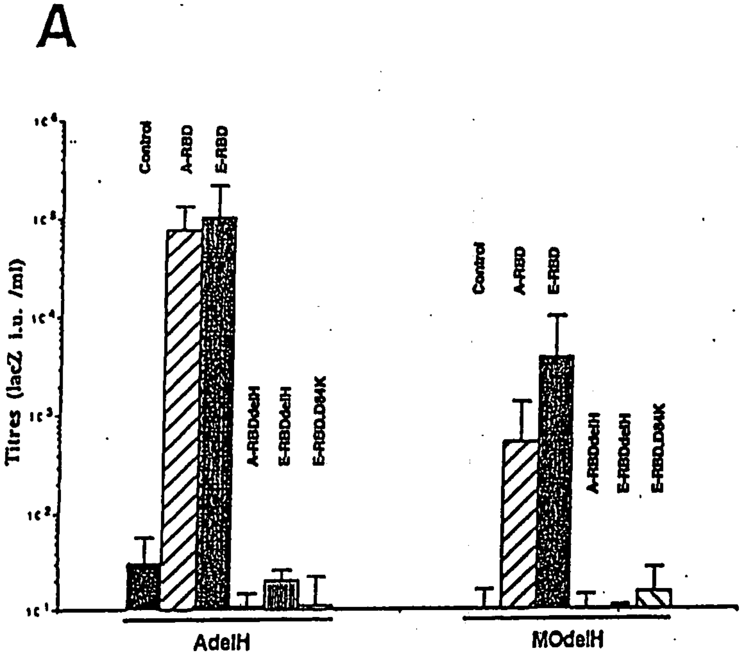
A
Titres (lacZ i.u. /ml)
10^1
10^2
10^3
10^4
10^5
10^6
Control
A-RBD
E-RBD
A-RBDdelH
E-RBDdelH
E-RBD.D84K
Control
A-RBD
E-RBD
A-RBDdelH
E-RBDdelH
E-RBD.D84K
AdelH
MOdelH

FIG.5A

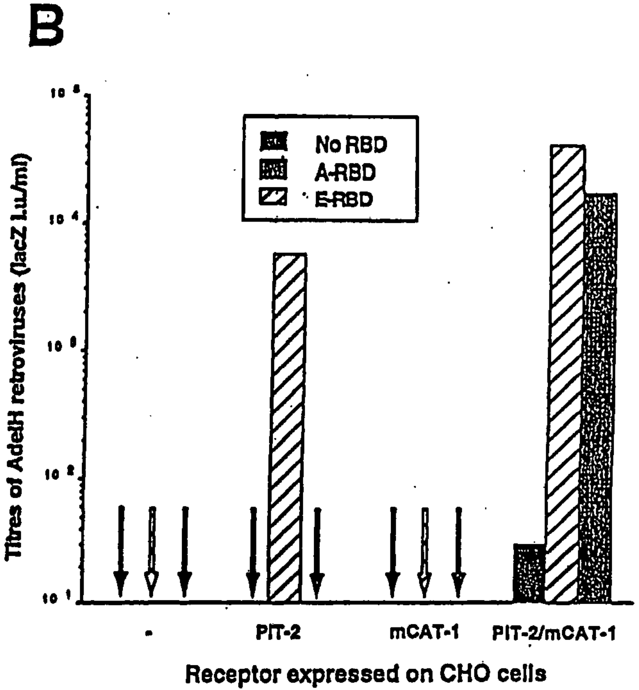
B
Titres of AdelH retroviruses (lacZ i.u/ml)
10^1
10^2
10^3
10^4
10^5
No RBD
A-RBD
E-RBD
-
PIT-2
mCAT-1
PIT-2/mCAT-1
Receptor expressed on CHO cells

FIG.5B

A-RBD Sequence

```
            10              20              30              40              50              60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
ATG GCG CGT TCA ACG CTC TCA AAA CCC CCT CAA GAT AAG ATT AAC CCG TGG AAG CCC TTA     60
Met Ala Arg Ser Thr Leu Ser Lys Pro Pro Gln Asp Lys Ile Asn Pro Trp Lys Pro Leu

ATA GTC ATG GGA GTC CTG TTA GGA GTA GGG ATG GCA GAG AGC CCC CAT CAG GTC TTT AAT    120
Ile Val Met Gly Val Leu Leu Gly Val Gly Met Ala Glu Ser Pro His Gln Val Phe Asn

GTA ACC TGG AGA GTC ACC AAC CTG ATG ACT GGG CGT ACC GCC AAT GCC ACC TCC CTC CTG    180
Val Thr Trp Arg Val Thr Asn Leu Met Thr Gly Arg Thr Ala Asn Ala Thr Ser Leu Leu

GGA ACT GTA CAA GAT GCC TTC CCA AAA TTA TAT TTT GAT CTA TGT GAT CTG GTC GGA GAG    240
Gly Thr Val Gln Asp Ala Phe Pro Lys Leu Tyr Phe Asp Leu Cys Asp Leu Val Gly Glu

GAG TGG GAC CCT TCA GAC CAG GAA CCG TAT GTC GGG TAT GGC TGC AAG TAC CCC GCA GGG    300
Glu Trp Asp Pro Ser Asp Gln Glu Pro Tyr Val Gly Tyr Gly Cys Lys Tyr Pro Ala Gly

AGA CAG CGG ACC CGG ACT TTT GAC TTT TAC GTG TGC CCT GGG CAT ACC GTA AAG TCG GGG    360
Arg Gln Arg Thr Arg Thr Phe Asp Phe Tyr Val Cys Pro Gly His Thr Val Lys Ser Gly

TGT GGG GGA CCA GGA GAG GGC TAC TGT GGT AAA TGG GGG TGT GAA ACC ACC GGA CAG GCT    420
Cys Gly Gly Pro Gly Glu Gly Tyr Cys Gly Lys Trp Gly Cys Glu Thr Thr Gly Gln Ala

TAC TGG AAG CCC ACA TCA TCG TGG GAC CTA ATC TCC CTT AAG CGC GGT AAC ACC CCC TGG    480
Tyr Trp Lys Pro Thr Ser Ser Trp Asp Leu Ile Ser Leu Lys Arg Gly Asn Thr Pro Trp

GAC ACG GGA TGC TCT AAA GTT GCC TGT GGC CCC TGC TAC GAC CTC TCC AAA GTA TCC AAT    540
Asp Thr Gly Cys Ser Lys Val Ala Cys Gly Pro Cys Tyr Asp Leu Ser Lys Val Ser Asn

TCC TTC CAA GGG GCT ACT CGA GGG GGC AGA TGC AAC CCT CTA GTC CTA GAA TTC ACT GAT    600
Ser Phe Gln Gly Ala Thr Arg Gly Gly Arg Cys Asn Pro Leu Val Leu Glu Phe Thr Asp

GCA GGA AAA AAG GCT AAC TGG GAC GGG CCC AAA TCG TGG GGA CTG AGA CTG TAC CGG ACA    660
Ala Gly Lys Lys Ala Asn Trp Asp Gly Pro Lys Ser Trp Gly Leu Arg Leu Tyr Arg Thr

GGA ACA GAT CCT ATT ACC ATG TTC TCC CTG ACC CGG CAG GTC CTT AAT GTG GGA CCC CGA    720
Gly Thr Asp Pro Ile Thr Met Phe Ser Leu Thr Arg Gln Val Leu Asn Val Gly Pro Arg

GTC CCC ATA GGG ATC CAA TAC CCC TAC GAC GTC CCC GAC TAT GCC AAT CGA TTA GTT CAA    780
Val Pro Ile Gly Ile Gln Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Asn Arg Leu Val Gln

TTT GTT AAA GAC AGG ATC TCA GTA GTC CAG GCT TTA GTC CTG ACT CAA CAA TAC CAC CAG    840
Phe Val Lys Asp Arg Ile Ser Val Val Gln Ala Leu Val Leu Thr Gln Gln Tyr His Gln

CTA AAA CCA CTA GAA TAC GAG CCA TGA                                                867
Leu Lys Pro Leu Glu Tyr Glu Pro ...
```

FIG. 6

E-RBD- Sequence

```
                10          20          30          40          50          60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
ATG GCG CGT TCA ACG CTC TCA AAA CCC CTT AAA AAT AAG GTT AAC CCG CGA GGC CCC CTA     60
Met Ala Arg Ser Thr Leu Ser Lys Pro Leu Lys Asn Lys Val Asn Pro Arg Gly Pro Leu

ATC CCC TTA ATT CTT CTG ATG CTC AGA GGG GTC AGT ACT GCT TCG CCC GGC TCC AGT CCT    120
Ile Pro Leu Ile Leu Leu Met Leu Arg Gly Val Ser Thr Ala Ser Pro Gly Ser Ser Pro

CAT CAA GTC TAT AAT ATC ACC TGG GAG GTA ACC AAT GGA GAT CGG GAG ACG GTA TGG GCA    180
His Gln Val Tyr Asn Ile Thr Trp Glu Val Thr Asn Gly Asp Arg Glu Thr Val Trp Ala

ACT TCT GGC AAC CAC CCT CTG TGG ACC TGG TGG CCT GAC CTT ACC CCA GAT TTA TGT ATG    240
Thr Ser Gly Asn His Pro Leu Trp Thr Trp Trp Pro Asp Leu Thr Pro Asp Leu Cys Met

TTA GCC CAC CAT GGA CCA TCT TAT TGG GGG CTA GAA TAT CAA TCC CCT TTT TCT TCT CCC    300
Leu Ala His His Gly Pro Ser Tyr Trp Gly Leu Glu Tyr Gln Ser Pro Phe Ser Ser Pro

CCG GGG CCC CCT TGT TGC TCA GGG GGC AGC AGC CCA GGC TGT TCC AGA GAC TGC GAA GAA    360
Pro Gly Pro Pro Cys Cys Ser Gly Gly Ser Ser Pro Gly Cys Ser Arg Asp Cys Glu Glu

CCT TTA ACC TCC CTC ACC CCT CGG TGC AAC ACT GCC TGG AAC AGA CTC AAG CTA GAC CAG    420
Pro Leu Thr Ser Leu Thr Pro Arg Cys Asn Thr Ala Trp Asn Arg Leu Lys Leu Asp Gln

ACA ACT CAT AAA TCA AAT GAG GGA TTT TAT GTT TGC CCC GGG CCC CAC CGC CCC CGA GAA    480
Thr Thr His Lys Ser Asn Glu Gly Phe Tyr Val Cys Pro Gly Pro His Arg Pro Arg Glu

TCC AAG TCA TGT GGG GGT CCA GAC TCC TTC TAC TGT GCC TAT TGG GGC TGT GAG ACA ACC    540
Ser Lys Ser Cys Gly Gly Pro Asp Ser Phe Tyr Cys Ala Tyr Trp Gly Cys Glu Thr Thr

GGT AGA GCT TAC TGG AAG CCC TCC TCA TCA TGG GAT TTC ATC ACA GTA AAC AAC AAT CTC    600
Gly Arg Ala Tyr Trp Lys Pro Ser Ser Ser Trp Asp Phe Ile Thr Val Asn Asn Asn Leu

ACC TCT GAC CAG GCT GTC CAG GTA TGC AAA GAT AAT AAG TGG TGC AAC CCC TTA GTT ATT    660
Thr Ser Asp Gln Ala Val Gln Val Cys Lys Asp Asn Lys Trp Cys Asn Pro Leu Val Ile

CGG TTT ACA GAC GCC GGG AGA CGG GTT ACT TCC TGG ACC ACA GGA CAT TAC TGG GGC TTA    720
Arg Phe Thr Asp Ala Gly Arg Arg Val Thr Ser Trp Thr Thr Gly His Tyr Trp Gly Leu

CGT TTG TAT GTC TCC GGA CAA GAT CCA GGG CTT ACA TTT GGG ATC AGA CTC AGA TAC CAA    780
Arg Leu Tyr Val Ser Gly Gln Asp Pro Gly Leu Thr Phe Gly Ile Arg Leu Arg Tyr Gln

AAT CTA GGA CCC CGC GTC CCA ATA GGG ATC CAA TAC CCC TAC GAC GTC CCC GAC TAT GCC    840
Asn Leu Gly Pro Arg Val Pro Ile Gly Ile Gln Tyr Pro Tyr Asp Val Pro Asp Tyr Ala

AAT CGA TTA GTT CAA TTT GTT AAA GAC AGG ATC TCA GTA GTC CAG GCT TTA GTC CTG ACT    900
Asn Arg Leu Val Gln Phe Val Lys Asp Arg Ile Ser Val Val Gln Ala Leu Val Leu Thr

CAA CAA TAC CAC CAG CTA AAA CCA CTA GAA TAC GAG CCA TGA
Gln Gln Tyr His Gln Leu Lys Pro Leu Glu Tyr Glu Pro ...                            942
```

FIG. 7

GALV-RBD- Sequence

```
            10          20          30          40          50          60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
ATG GTA TTG CTG CCT GGG TCC ATG CTT CTC ACC TCA AAC CTG CAC CAC CTT CGG CAC CAG    60
Met Val Leu Leu Pro Gly Ser Met Leu Leu Thr Ser Asn Leu His His Leu Arg His Gln

ATG AGT CCT GGG AGC TGG AAA AGA CTG ATC ATC CTC TTA AGC TGC GTA TTC GGC GGC GGC   120
Met Ser Pro Gly Ser Trp Lys Arg Leu Ile Ile Leu Leu Ser Cys Val Phe Gly Gly Gly

GGG ACG AGT CTG CAA AAT AAG AAC CCC CAC CAG CCC ATG ACC CTC ACT TGG CAG GTA CTG   180
Gly Thr Ser Leu Gln Asn Lys Asn Pro His Gln Pro Met Thr Leu Thr Trp Gln Val Leu

TCC CAA ACT GGA GAC GTT GTC TGG GAT ACA AAG GCA GTC CAG CCC CCT TGG ACT TGG TGG   240
Ser Gln Thr Gly Asp Val Val Trp Asp Thr Lys Ala Val Gln Pro Pro Trp Thr Trp Trp

CCC ACA CTT AAA CCT GAT GTA TGT GCC TTG GCG GCT AGT CTT GAG TCC TGG GAT ATC CCG   300
Pro Thr Leu Lys Pro Asp Val Cys Ala Leu Ala Ala Ser Leu Glu Ser Trp Asp Ile Pro

GGA ACC GAT GTC TCG TCC TCT AAA CGA GTC AGA CCT CCG GAC TCA GAC TAT ACT GCC GCT   360
Gly Thr Asp Val Ser Ser Ser Lys Arg Val Arg Pro Pro Asp Ser Asp Tyr Thr Ala Ala

TAT AAG CAA ATC ACC TGG GGA GCC ATA GGG TGC AGC TAC CCT CGG GCT AGG ACT AGA ATG   420
Tyr Lys Gln Ile Thr Trp Gly Ala Ile Gly Cys Ser Tyr Pro Arg Ala Arg Thr Arg Met

GCA AGC TCT ACC TTC TAC GTA TGT CCC CGG GAT GGC CGG ACC CTT TCA GAA GCT AGA AGG   480
Ala Ser Ser Thr Phe Tyr Val Cys Pro Arg Asp Gly Arg Thr Leu Ser Glu Ala Arg Arg

TGC GGG GGG CTA GAA TCC CTA TAC TGT AAA GAA TGG GAT TGT GAG ACC ACG GGG ACC GGT   540
Cys Gly Gly Leu Glu Ser Leu Tyr Cys Lys Glu Trp Asp Cys Glu Thr Thr Gly Thr Gly

TAT TGG CTA TCT AAA TCC TCA AAA GAC CTC ATA ACT GTA AAA TGG GAC CAA AAT AGC GAA   600
Tyr Trp Leu Ser Lys Ser Ser Lys Asp Leu Ile Thr Val Lys Trp Asp Gln Asn Ser Glu

TGG ACT CAA AAA TTT CAA CAG TGT CAC CAG ACC GGC TGG TGT AAC CCC CTT AAA ATA GAT   660
Trp Thr Gln Lys Phe Gln Gln Cys His Gln Thr Gly Trp Cys Asn Pro Leu Lys Ile Asp

TTC ACA GAC AAA GGA AAA TTA TCC AAG GAC TGG ATA ACG GGA AAA ACC TGG GGA TTA AGA   720
Phe Thr Asp Lys Gly Lys Leu Ser Lys Asp Trp Ile Thr Gly Lys Thr Trp Gly Leu Arg

TTC TAT GTG TCT GGA CAT CCA GGC GTA CAG TTC ACC ATT CGC TTA AAA ATC ACC AAC ATG   780
Phe Tyr Val Ser Gly His Pro Gly Val Gln Phe Thr Ile Arg Leu Lys Ile Thr Asn Met

CCA GCT GTG GCA GTA GGG ATC CAA TAC CCC TAC GAC GTC CCC GAC TAT GCC AAT CGA TTA   840
Pro Ala Val Ala Val Gly Ile Gln Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Asn Arg Leu

GTT CAA TTT GTT AAA GAC AGG ATC TCA GTA GTC CAG GCT TTA GTC CTG ACT CAA CAA TAC   900
Val Gln Phe Val Lys Asp Arg Ile Ser Val Val Gln Ala Leu Val Leu Thr Gln Gln Tyr

CAC CAG CTA AAA CCA CTA GAA TAC GAG CCA TGA                                       933
His Gln Leu Lys Pro Leu Glu Tyr Glu Pro ...
```

FIG. 8

EGF-A-RBD- Sequence

```
            10          20          30          40          50          60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
ATG GCG CGT TCA ACG CTC TCA AAA CCC CTT AAA AAT AAG GTT AAC CCG CGA GGC CCC CTA    60
Met Ala Arg Ser Thr Leu Ser Lys Pro Leu Lys Asn Lys Val Asn Pro Arg Gly Pro Leu

ATC CCC TTA ATT CTT CTG ATG CTC AGA GGG GTC AGT ACT GCT TCG CCC GGC TCC AGT GCG   120
Ile Pro Leu Ile Leu Leu Met Leu Arg Gly Val Ser Thr Ala Ser Pro Gly Ser Ser Ala

GCC CAG CCG GCC ATG GCC AAT AGT GAC TCT GAA TGT CCC CTG TCC CAC GAT GGG TAC TGC   180
Ala Gln Pro Ala Met Ala Asn Ser Asp Ser Glu Cys Pro Leu Ser His Asp Gly Tyr Cys

CTC CAT GAT GGT GTG TGC ATG TAT ATT GAA GCA TTG GAC AAG TAT GCA TGC AAC TGT GTT   240
Leu His Asp Gly Val Cys Met Tyr Ile Glu Ala Leu Asp Lys Tyr Ala Cys Asn Cys Val

GTT GGC TAC ATC GGG GAG CGA TGT CAG TAC CGA GAC CTG AAG TGG TGG GAA CTG CGC GCG   300
Val Gly Tyr Ile Gly Glu Arg Cys Gln Tyr Arg Asp Leu Lys Trp Trp Glu Leu Arg Ala

GCC GCA ATC GAG GGA AGG ATG GCA GAG AGC CCC CAT CAG GTC TTT AAT GTA ACC TGG AGA   360
Ala Ala Ile Glu Gly Arg Met Ala Glu Ser Pro His Gln Val Phe Asn Val Thr Trp Arg

GTC ACC AAC CTG ATG ACT GGG CGT ACC GCC AAT GCC ACC TCC CTC CTG GGA ACT GTA CAA   420
Val Thr Asn Leu Met Thr Gly Arg Thr Ala Asn Ala Thr Ser Leu Leu Gly Thr Val Gln

GAT GCC TTC CCA AAA TTA TAT TTT GAT CTA TGT GAT CTG GTC GGA GAG GAG TGG GAC CCT   480
Asp Ala Phe Pro Lys Leu Tyr Phe Asp Leu Cys Asp Leu Val Gly Glu Glu Trp Asp Pro

TCA GAC CAG GAA CCG TAT GTC GGG TAT GGC TGC AAG TAC CCC GCA GGG AGA CAG CGG ACC   540
Ser Asp Gln Glu Pro Tyr Val Gly Tyr Gly Cys Lys Tyr Pro Ala Gly Arg Gln Arg Thr

CGG ACT TTT GAC TTT TAC GTG TGC CCT GGG CAT ACC GTA AAG TCG GGG TGT GGG GGA CCA   600
Arg Thr Phe Asp Phe Tyr Val Cys Pro Gly His Thr Val Lys Ser Gly Cys Gly Gly Pro

GGA GAG GGC TAC TGT GGT AAA TGG GGG TGT GAA ACC ACC GGA CAG GCT TAC TGG AAG CCC   660
Gly Glu Gly Tyr Cys Gly Lys Trp Gly Cys Glu Thr Thr Gly Gln Ala Tyr Trp Lys Pro

ACA TCA TCG TGG GAC CTA ATC TCC CTT AAG CGC GGT AAC ACC CCC TGG GAC ACG GGA TGC   720
Thr Ser Ser Trp Asp Leu Ile Ser Leu Lys Arg Gly Asn Thr Pro Trp Asp Thr Gly Cys

TCT AAA GTT GCC TGT GGC CCC TGC TAC GAC CTC TCC AAA GTA TCC AAT TCC TTC CAA GGG   780
Ser Lys Val Ala Cys Gly Pro Cys Tyr Asp Leu Ser Lys Val Ser Asn Ser Phe Gln Gly

GCT ACT CGA GGG GGC AGA TGC AAC CCT CTA GTC CTA GAA TTC ACT GAT GCA GGA AAA AAG   840
Ala Thr Arg Gly Gly Arg Cys Asn Pro Leu Val Leu Glu Phe Thr Asp Ala Gly Lys Lys

GCT AAC TGG GAC GGG CCC AAA TCG TGG GGA CTG AGA CTG TAC CGG ACA GGA ACA GAT CCT   900
Ala Asn Trp Asp Gly Pro Lys Ser Trp Gly Leu Arg Leu Tyr Arg Thr Gly Thr Asp Pro

ATT ACC ATG TTC TCC CTG ACC CGG CAG GTC CTT AAT GTG GGA CCC CGA GTC CCC ATA GGG   960
Ile Thr Met Phe Ser Leu Thr Arg Gln Val Leu Asn Val Gly Pro Arg Val Pro Ile Gly

ATC CAA TAC CCC TAC GAC GTC CCC GAC TAT GCC AAT CGA TTA GTT CAA TTT GTT AAA GAC  1020
Ile Gln Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Asn Arg Leu Val Gln Phe Val Lys Asp

AGG ATC TCA GTA GTC CAG GCT TTA GTC CTG ACT CAA CAA TAC CAC CAG CTA AAA CCA CTA  1080
Arg Ile Ser Val Val Gln Ala Leu Val Leu Thr Gln Gln Tyr His Gln Leu Lys Pro Leu

GAA TAC GAG CCA TGA
Glu Tyr Glu Pro ...                                                              1095
```

FIG. 9

EGF-E-RBD- Sequence

```
               10                  20                  30                  40                  50                  60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
ATG GCG CGT TCA ACG CTC TCA AAA CCC CTT AAA AAT AAG GTT AAC CCG CGA GGC CCC CTA      60
Met Ala Arg Ser Thr Leu Ser Lys Pro Leu Lys Asn Lys Val Asn Pro Arg Gly Pro Leu

ATC CCC TTA ATT CTT CTG ATG CTC AGA GGG GTC AGT ACT GCT TCG CCC GGC TCC AGT GCG     120
Ile Pro Leu Ile Leu Leu Met Leu Arg Gly Val Ser Thr Ala Ser Pro Gly Ser Ser Ala

GCC CAG CCG GCC ATG GCC AAT AGT GAC TCT GAA TGT CCC CTG TCC CAC GAT GGG TAC TGC     180
Ala Gln Pro Ala Met Ala Asn Ser Asp Ser Glu Cys Pro Leu Ser His Asp Gly Tyr Cys

CTC CAT GAT GGT GTG TGC ATG TAT ATT GAA GCA TTG GAC AAG TAT GCA TGC AAC TGT GTT     240
Leu His Asp Gly Val Cys Met Tyr Ile Glu Ala Leu Asp Lys Tyr Ala Cys Asn Cys Val

GTT GGC TAC ATC GGG GAG CGA TGT CAG TAC CGA GAC CTG AAG TGG TGG GAA CTG CGC GCG     300
Val Gly Tyr Ile Gly Glu Arg Cys Gln Tyr Arg Asp Leu Lys Trp Trp Glu Leu Arg Ala

GCC GCA ATC GAG GGA AGG GCT TCG CCC GGC TCC AGT CCT CAT CAA GTC TAT AAT ATC ACC     360
Ala Ala Ile Glu Gly Arg Ala Ser Pro Gly Ser Ser Pro His Gln Val Tyr Asn Ile Thr

TGG GAG GTA ACC AAT GGA GAT CGG GAG ACG GTA TGG GCA ACT TCT GGC AAC CAC CCT CTG     420
Trp Glu Val Thr Asn Gly Asp Arg Glu Thr Val Trp Ala Thr Ser Gly Asn His Pro Leu

TGG ACC TGG TGG CCT GAC CTT ACC CCA GAT TTA TGT ATG TTA GCC CAC CAT GGA CCA TCT     480
Trp Thr Trp Trp Pro Asp Leu Thr Pro Asp Leu Cys Met Leu Ala His His Gly Pro Ser

TAT TGG GGG CTA GAA TAT CAA TCC CCT TTT TCT TCT CCC CCG GGG CCC CCT TGT TGC TCA     540
Tyr Trp Gly Leu Glu Tyr Gln Ser Pro Phe Ser Ser Pro Pro Gly Pro Pro Cys Cys Ser

GGG GGC AGC AGC CCA GGC TGT TCC AGA GAC TGC GAA GAA CCT TTA ACC TCC CTC ACC CCT     600
Gly Gly Ser Ser Pro Gly Cys Ser Arg Asp Cys Glu Glu Pro Leu Thr Ser Leu Thr Pro

CGG TGC AAC ACT GCC TGG AAC AGA CTC AAG CTA GAC CAG ACA ACT CAT AAA TCA AAT GAG     660
Arg Cys Asn Thr Ala Trp Asn Arg Leu Lys Leu Asp Gln Thr Thr His Lys Ser Asn Glu

GGA TTT TAT GTT TGC CCC GGG CCC CAC CGC CCC CGA GAA TCC AAG TCA TGT GGG GGT CCA     720
Gly Phe Tyr Val Cys Pro Gly Pro His Arg Pro Arg Glu Ser Lys Ser Cys Gly Gly Pro

GAC TCC TTC TAC TGT GCC TAT TGG GGC TGT GAG ACA ACC GGT AGA GCT TAC TGG AAG CCC     780
Asp Ser Phe Tyr Cys Ala Tyr Trp Gly Cys Glu Thr Thr Gly Arg Ala Tyr Trp Lys Pro

TCC TCA TCA TGG GAT TTC ATC ACA GTA AAC AAC AAT CTC ACC TCT GAC CAG GCT GTC CAG     840
Ser Ser Ser Trp Asp Phe Ile Thr Val Asn Asn Asn Leu Thr Ser Asp Gln Ala Val Gln

GTA TGC AAA GAT AAT AAG TGG TGC AAC CCC TTA GTT ATT CGG TTT ACA GAC GCC GGG AGA     900
Val Cys Lys Asp Asn Lys Trp Cys Asn Pro Leu Val Ile Arg Phe Thr Asp Ala Gly Arg

CGG GTT ACT TCC TGG ACC ACA GGA CAT TAC TGG GGC TTA CGT TTG TAT GTC TCC GGA CAA     960
Arg Val Thr Ser Trp Thr Thr Gly His Tyr Trp Gly Leu Arg Leu Tyr Val Ser Gly Gln

GAT CCA GGG CTT ACA TTT GGG ATC AGA CTC AGA TAC CAA AAT CTA GGA CCC CGC GTC CCA    1020
Asp Pro Gly Leu Thr Phe Gly Ile Arg Leu Arg Tyr Gln Asn Leu Gly Pro Arg Val Pro

ATA GGG ATC CAA TAC CCC TAC GAC GTC CCC GAC TAT GCC AAT CGA TTA GTT CAA TTT GTT    1080
Ile Gly Ile Gln Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Asn Arg Leu Val Gln Phe Val

AAA GAC AGG ATC TCA GTA GTC CAG GCT TTA GTC CTG ACT CAA CAA TAC CAC CAG CTA AAA    1140
Lys Asp Arg Ile Ser Val Val Gln Ala Leu Val Leu Thr Gln Gln Tyr His Gln Leu Lys

CCA CTA GAA TAC GAG CCA TGA
Pro Leu Glu Tyr Glu Pro ...                                                         1161
```

FIG. 10

EGF-GALV-RBD- . Sequence

```
                   10                  20                  30                  40                  50                  60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
ATG GCG CGT TCA ACG CTC TCA AAA CCC CTT AAA AAT AAG GTT AAC CCG CGA GGC CCC CTA     60
Met Ala Arg Ser Thr Leu Ser Lys Pro Leu Lys Asn Lys Val Asn Pro Arg Gly Pro Leu

ATC CCC TTA ATT CTT CTG ATG CTC AGA GGG GTC AGT ACT GCT TCG CCC GGC TCC AGT GCG    120
Ile Pro Leu Ile Leu Leu Met Leu Arg Gly Val Ser Thr Ala Ser Pro Gly Ser Ser Ala

GCC CAG CCG GCC ATG GCC AAT AGT GAC TCT GAA TGT CCC CTG TCC CAC GAT GGG TAC TGC    180
Ala Gln Pro Ala Met Ala Asn Ser Asp Ser Glu Cys Pro Leu Ser His Asp Gly Tyr Cys

CTC CAT GAT GGT GTG TGC ATG TAT ATT GAA GCA TTG GAC AAG TAT GCA TGC AAC TGT GTT    240
Leu His Asp Gly Val Cys Met Tyr Ile Glu Ala Leu Asp Lys Tyr Ala Cys Asn Cys Val

GTT GGC TAC ATC GGG GAG CGA TGT CAG TAC CGA GAC CTG AAG TGG TGG GAA CTG CGC GCG    300
Val Gly Tyr Ile Gly Glu Arg Cys Gln Tyr Arg Asp Leu Lys Trp Trp Glu Leu Arg Ala

GCC GCA ATC GAG GGA AGG AGT CTG CAA AAT AAG AAC CCC CAC CAG CCC ATG ACC CTC ACT    360
Ala Ala Ile Glu Gly Arg Ser Leu Gln Asn Lys Asn Pro His Gln Pro Met Thr Leu Thr

TGG CAG GTA CTG TCC CAA ACT GGA GAC GTT GTC TGG GAT ACA AAG GCA GTC CAG CCC CCT    420
Trp Gln Val Leu Ser Gln Thr Gly Asp Val Val Trp Asp Thr Lys Ala Val Gln Pro Pro

TGG ACT TGG TGG CCC ACA CTT AAA CCT GAT GTA TGT GCC TTG GCG GCT AGT CTT GAG TCC    480
Trp Thr Trp Trp Pro Thr Leu Lys Pro Asp Val Cys Ala Leu Ala Ala Ser Leu Glu Ser

TGG GAT ATC CCG GGA ACC GAT GTC TCG TCC TCT AAA CGA GTC AGA CCT CCG GAC TCA GAC    540
Trp Asp Ile Pro Gly Thr Asp Val Ser Ser Ser Lys Arg Val Arg Pro Pro Asp Ser Asp

TAT ACT GCC GCT TAT AAG CAA ATC ACC TGG GGA GCC ATA GGG TGC AGC TAC CCT CGG GCT    600
Tyr Thr Ala Ala Tyr Lys Gln Ile Thr Trp Gly Ala Ile Gly Cys Ser Tyr Pro Arg Ala

AGG ACT AGA ATG GCA AGC TCT ACC TTC TAC GTA TGT CCC CGG GAT GGC CGG ACC CTT TCA    660
Arg Thr Arg Met Ala Ser Ser Thr Phe Tyr Val Cys Pro Arg Asp Gly Arg Thr Leu Ser

GAA GCT AGA AGG TGC GGG GGG CTA GAA TCC CTA TAC TGT AAA GAA TGG GAT TGT GAG ACC    720
Glu Ala Arg Arg Cys Gly Gly Leu Glu Ser Leu Tyr Cys Lys Glu Trp Asp Cys Glu Thr

ACG GGG ACC GGT TAT TGG CTA TCT AAA TCC TCA AAA GAC CTC ATA ACT GTA AAA TGG GAC    780
Thr Gly Thr Gly Tyr Trp Leu Ser Lys Ser Ser Lys Asp Leu Ile Thr Val Lys Trp Asp

CAA AAT AGC GAA TGG ACT CAA AAA TTT CAA CAG TGT CAC CAG ACC GGC TGG TGT AAC CCC    840
Gln Asn Ser Glu Trp Thr Gln Lys Phe Gln Gln Cys His Gln Thr Gly Trp Cys Asn Pro

CTT AAA ATA GAT TTC ACA GAC AAA GGA AAA TTA TCC AAG GAC TGG ATA ACG GGA AAA ACC    900
Leu Lys Ile Asp Phe Thr Asp Lys Gly Lys Leu Ser Lys Asp Trp Ile Thr Gly Lys Thr

TGG GGA TTA AGA TTC TAT GTG TCT GGA CAT CCA GGC GTA CAG TTC ACC ATT CGC TTA AAA    960
Trp Gly Leu Arg Phe Tyr Val Ser Gly His Pro Gly Val Gln Phe Thr Ile Arg Leu Lys

ATC ACC AAC ATG CCA GCT GTG GCA GTA GGG ATC CAA TAC CCC TAC GAC GTC CCC GAC TAT   1020
Ile Thr Asn Met Pro Ala Val Ala Val Gly Ile Gln Tyr Pro Tyr Asp Val Pro Asp Tyr

GCC AAT CGA TTA GTT CAA TTT GTT AAA GAC AGG ATC TCA GTA GTC CAG GCT TTA GTC CTG   1080
Ala Asn Arg Leu Val Gln Phe Val Lys Asp Arg Ile Ser Val Val Gln Ala Leu Val Leu

ACT CAA CAA TAC CAC CAG CTA AAA CCA CTA GAA TAC GAG CCA TGA                       1125
Thr Gln Gln Tyr His Gln Leu Lys Pro Leu Glu Tyr Glu Pro ...
```

FIG. 11

34.1-A-RBD- Sequence

```
            10          20          30          40          50          60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
ATG GCG CGT TCA ACG CTC TCA AAA CCC CTT AAA AAT AAG GTT AAC CCG CGA GGC CCC CTA     60
Met Ala Arg Ser Thr Leu Ser Lys Pro Leu Lys Asn Lys Val Asn Pro Arg Gly Pro Leu

ATC CCC TTA ATT CTT CTG ATG CTC AGA GGG GTC AGT ACT GCT TCG CCC GGC TCC AGT GCG    120
Ile Pro Leu Ile Leu Leu Met Leu Arg Gly Val Ser Thr Ala Ser Pro Gly Ser Ser Ala

GCC CAG CCG GCC ATG GCC CAG GTT CAG CTT CAG CAG TCT GGA GCT GAA CTG GTA AAG CCT    180
Ala Gln Pro Ala Met Ala Gln Val Gln Leu Gln Gln Ser Gly Ala Glu Leu Val Lys Pro

GGG GCT TCA GTG AAG TTG TCC TGC AAG GCT TCT GGC TAC ACC TTC ACA AGC TAT GAT ATA    240
Gly Ala Ser Val Lys Leu Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr Asp Ile

AAC TGG GTG AGG CAG AGG CCT GAA CAG GGA CTT GAG TGG ATT GGA TGG ATT TTT CCT GGA    300
Asn Trp Val Arg Gln Arg Pro Glu Gln Gly Leu Glu Trp Ile Gly Trp Ile Phe Pro Gly

GAT GGT AGT ACT AAG TAC AAT GAG AAG TTC AAG GGC AAG GCC ACA CTG ACT ACA GAC AAA    360
Asp Gly Ser Thr Lys Tyr Asn Glu Lys Phe Lys Gly Lys Ala Thr Leu Thr Thr Asp Lys

TCC TCC AGC ACA GCC TAC ATG CAG CTC AGC AGG CTG ACA TCT GAA GAC TCT GCT ATC TAT    420
Ser Ser Ser Thr Ala Tyr Met Gln Leu Ser Arg Leu Thr Ser Glu Asp Ser Ala Ile Tyr

TTC TGT GCA AGA ATG GCC GAC TAT GGT CAC TAC GGC GCT GTG GAC TAC TGG GGT CAA GGG    480
Phe Cys Ala Arg Met Ala Asp Tyr Gly His Tyr Gly Ala Val Asp Tyr Trp Gly Gln Gly

ACC ACG GTC ACC GTC TCC TCA GGT GGA GGC GGT TCA GGC GGA GGT GGC TCT GGC GGT GGC    540
Thr Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly

GGA TCG GAC ATC CAG ATG ACT CAG TCT CCA TCC TCC CTG TCT GCC TCT CTG GGA GAC AGA    600
Gly Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Leu Gly Asp Arg

GTC ACC ATC AGT TGC AGG GCA AGT CAG GAC ATT AGC AAT TAT TTA AAC TGG TAT CAG CAG    660
Val Thr Ile Ser Cys Arg Ala Ser Gln Asp Ile Ser Asn Tyr Leu Asn Trp Tyr Gln Gln

ACA CCA GAT GGA ACT GTT AAA CTC CTG ATC TAC TAC ACA TCA AGA TTA CAC TCA GGA GTC    720
Thr Pro Asp Gly Thr Val Lys Leu Leu Ile Tyr Tyr Thr Ser Arg Leu His Ser Gly Val

CCA TCA AGG TTC AGT GGC AGT GGG TCT GGA ACA GAT TAT TCT CTC ACC ATT AGC AAC CTG    780
Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Tyr Ser Leu Thr Ile Ser Asn Leu

GAG CAA GAA GAT ATT GCC ACT TAC TTT TGC CAA CAG GGT TAT ACG CTT CCG TGG ACG TTC    840
Glu Gln Glu Asp Ile Ala Thr Tyr Phe Cys Gln Gln Gly Tyr Thr Leu Pro Trp Thr Phe

GGT GGA GGC ACC AAG CTG GAG ATC AAA CGG GCG GCC GCA ATC GAG GGA AGG ATG GCA GAG    900
Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg Ala Ala Ala Ile Glu Gly Arg Met Ala Glu

AGC CCC CAT CAG GTC TTT AAT GTA ACC TGG AGA GTC ACC AAC CTG ATG ACT GGG CGT ACC    960
Ser Pro His Gln Val Phe Asn Val Thr Trp Arg Val Thr Asn Leu Met Thr Gly Arg Thr

GCC AAT GCC ACC TCC CTC CTG GGA ACT GTA CAA GAT GCC TTC CCA AAA TTA TAT TTT GAT   1020
Ala Asn Ala Thr Ser Leu Leu Gly Thr Val Gln Asp Ala Phe Pro Lys Leu Tyr Phe Asp

CTA TGT GAT CTG GTC GGA GAG GAG TGG GAC CCT TCA GAC CAG GAA CCG TAT GTC GGG TAT   1080
Leu Cys Asp Leu Val Gly Glu Glu Trp Asp Pro Ser Asp Gln Glu Pro Tyr Val Gly Tyr
```

FIG. 12

34.1-A-RBD- Sequence

```
               10                  20                  30                  40                  50                  60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890

GGC TGC AAG TAC CCC GCA GGG AGA CAG CGG ACC CGG ACT TTT GAC TTT TAC GTG TGC CCT   1140
Gly Cys Lys Tyr Pro Ala Gly Arg Gln Arg Thr Arg Thr Phe Asp Phe Tyr Val Cys Pro

GGG CAT ACC GTA AAG TCG GGG TGT GGG GGA CCA GGA GAG GGC TAC TGT GGT AAA TGG GGG   1200
Gly His Thr Val Lys Ser Gly Cys Gly Gly Pro Gly Glu Gly Tyr Cys Gly Lys Trp Gly

TGT GAA ACC ACC GGA CAG GCT TAC TGG AAG CCC ACA TCA TCG TGG GAC CTA ATC TCC CTT   1260
Cys Glu Thr Thr Gly Gln Ala Tyr Trp Lys Pro Thr Ser Ser Trp Asp Leu Ile Ser Leu

AAG CGC GGT AAC ACC CCC TGG GAC ACG GGA TGC TCT AAA GTT GCC TGT GGC CCC TGC TAC   1320
Lys Arg Gly Asn Thr Pro Trp Asp Thr Gly Cys Ser Lys Val Ala Cys Gly Pro Cys Tyr

GAC CTC TCC AAA GTA TCC AAT TCC TTC CAA GGG GCT ACT CGA GGG GGC AGA TGC AAC CCT   1380
Asp Leu Ser Lys Val Ser Asn Ser Phe Gln Gly Ala Thr Arg Gly Gly Arg Cys Asn Pro

CTA GTC CTA GAA TTC ACT GAT GCA GGA AAA AAG GCT AAC TGG GAC GGG CCC AAA TCG TGG   1440
Leu Val Leu Glu Phe Thr Asp Ala Gly Lys Lys Ala Asn Trp Asp Gly Pro Lys Ser Trp

GGA CTG AGA CTG TAC CGG ACA GGA ACA GAT CCT ATT ACC ATG TTC TCC CTG ACC CGG CAG   1500
Gly Leu Arg Leu Tyr Arg Thr Gly Thr Asp Pro Ile Thr Met Phe Ser Leu Thr Arg Gln

GTC CTT AAT GTG GGA CCC CGA GTC CCC ATA GGG ATC CAA TAC CCC TAC GAC GTC CCC GAC   1560
Val Leu Asn Val Gly Pro Arg Val Pro Ile Gly Ile Gln Tyr Pro Tyr Asp Val Pro Asp

TAT GCC AAT CGA TTA GTT CAA TTT GTT AAA GAC AGG ATC TCA GTA GTC CAG GCT TTA GTC   1620
Tyr Ala Asn Arg Leu Val Gln Phe Val Lys Asp Arg Ile Ser Val Val Gln Ala Leu Val

CTG ACT CAA CAA TAC CAC CAG CTA AAA CCA CTA GAA TAC GAG CCA TGA                   1668
Leu Thr Gln Gln Tyr His Gln Leu Lys Pro Leu Glu Tyr Glu Pro ...
```

FIG. 12 (continued)

34.1-E-RBD- Sequence

```
                 10                20               30                 40                50                60
 123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
ATG GCG CGT TCA ACG CTC TCA AAA CCC CTT AAA AAT AAG GTT AAC CCG CGA GGC CCC CTA      60
Met Ala Arg Ser Thr Leu Ser Lys Pro Leu Lys Asn Lys Val Asn Pro Arg Gly Pro Leu

ATC CCC TTA ATT CTT CTG ATG CTC AGA GGG GTC AGT ACT GCT TCG CCC GGC TCC AGT GCG     120
Ile Pro Leu Ile Leu Leu Met Leu Arg Gly Val Ser Thr Ala Ser Pro Gly Ser Ser Ala

GCC CAG CCG GCC ATG GCC CAG GTT CAG CTT CAG CAG TCT GGA GCT GAA CTG GTA AAG CCT     180
Ala Gln Pro Ala Met Ala Gln Val Gln Leu Gln Gln Ser Gly Ala Glu Leu Val Lys Pro

GGG GCT TCA GTG AAG TTG TCC TGC AAG GCT TCT GGC TAC ACC TTC ACA AGC TAT GAT ATA     240
Gly Ala Ser Val Lys Leu Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr Asp Ile

AAC TGG GTG AGG CAG AGG CCT GAA CAG GGA CTT GAG TGG ATT GGA TGG ATT TTT CCT GGA     300
Asn Trp Val Arg Gln Arg Pro Glu Gln Gly Leu Glu Trp Ile Gly Trp Ile Phe Pro Gly

GAT GGT AGT ACT AAG TAC AAT GAG AAG TTC AAG GGC AAG GCC ACA CTG ACT ACA GAC AAA     360
Asp Gly Ser Thr Lys Tyr Asn Glu Lys Phe Lys Gly Lys Ala Thr Leu Thr Thr Asp Lys

TCC TCC AGC ACA GCC TAC ATG CAG CTC AGC AGG CTG ACA TCT GAA GAC TCT GCT ATC TAT     420
Ser Ser Ser Thr Ala Tyr Met Gln Leu Ser Arg Leu Thr Ser Glu Asp Ser Ala Ile Tyr

TTC TGT GCA AGA ATG GCC GAC TAT GGT CAC TAC GGC GCT GTG GAC TAC TGG GGT CAA GGG     480
Phe Cys Ala Arg Met Ala Asp Tyr Gly His Tyr Gly Ala Val Asp Tyr Trp Gly Gln Gly

ACC ACG GTC ACC GTC TCC TCA GGT GGA GGC GGT TCA GGC GGA GGT GGC TCT GGC GGT GGC     540
Thr Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly

GGA TCG GAC ATC CAG ATG ACT CAG TCT CCA TCC TCC CTG TCT GCC TCT CTG GGA GAC AGA     600
Gly Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Leu Gly Asp Arg

GTC ACC ATC AGT TGC AGG GCA AGT CAG GAC ATT AGC AAT TAT TTA AAC TGG TAT CAG CAG     660
Val Thr Ile Ser Cys Arg Ala Ser Gln Asp Ile Ser Asn Tyr Leu Asn Trp Tyr Gln Gln

ACA CCA GAT GGA ACT GTT AAA CTC CTG ATC TAC TAC ACA TCA AGA TTA CAC TCA GGA GTC     720
Thr Pro Asp Gly Thr Val Lys Leu Leu Ile Tyr Tyr Thr Ser Arg Leu His Ser Gly Val

CCA TCA AGG TTC AGT GGC AGT GGG TCT GGA ACA GAT TAT TCT CTC ACC ATT AGC AAC CTG     780
Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Tyr Ser Leu Thr Ile Ser Asn Leu

GAG CAA GAA GAT ATT GCC ACT TAC TTT TGC CAA CAG GGT TAT ACG CTT CCG TGG ACG TTC     840
Glu Gln Glu Asp Ile Ala Thr Tyr Phe Cys Gln Gln Gly Tyr Thr Leu Pro Trp Thr Phe

GGT GGA GGC ACC AAG CTG GAG ATC AAA CGG GCG GCC GCA ATC GAG GGA AGG GCT TCG CCC     900
Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg Ala Ala Ala Ile Glu Gly Arg Ala Ser Pro

GGC TCC AGT CCT CAT CAA GTC TAT AAT ATC ACC TGG GAG GTA ACC AAT GGA GAT CGG GAG     960
Gly Ser Ser Pro His Gln Val Tyr Asn Ile Thr Trp Glu Val Thr Asn Gly Asp Arg Glu

ACG GTA TGG GCA ACT TCT GGC AAC CAC CCT CTG TGG ACC TGG TGG CCT GAC CTT ACC CCA    1020
Thr Val Trp Ala Thr Ser Gly Asn His Pro Leu Trp Thr Trp Trp Pro Asp Leu Thr Pro

GAT TTA TGT ATG TTA GCC CAC CAT GGA CCA TCT TAT TGG GGG CTA GAA TAT CAA TCC CCT    1080
Asp Leu Cys Met Leu Ala His His Gly Pro Ser Tyr Trp Gly Leu Glu Tyr Gln Ser Pro
```

FIG. 13

34.1-E-RBD- Sequence

```
                10          20          30          40          50          60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
TTT TCT TCT CCC CCG GGG CCC CCT TGT TGC TCA GGG GGC AGC AGC CCA GGC TGT TCC AGA   1140
Phe Ser Ser Pro Pro Gly Pro Pro Cys Cys Ser Gly Gly Ser Ser Pro Gly Cys Ser Arg

GAC TGC GAA GAA CCT TTA ACC TCC CTC ACC CCT CGG TGC AAC ACT GCC TGG AAC AGA CTC   1200
Asp Cys Glu Glu Pro Leu Thr Ser Leu Thr Pro Arg Cys Asn Thr Ala Trp Asn Arg Leu

AAG CTA GAC CAG ACA ACT CAT AAA TCA AAT GAG GGA TTT TAT GTT TGC CCC GGG CCC CAC   1260
Lys Leu Asp Gln Thr Thr His Lys Ser Asn Glu Gly Phe Tyr Val Cys Pro Gly Pro His

CGC CCC CGA GAA TCC AAG TCA TGT GGG GGT CCA GAC TCC TTC TAC TGT GCC TAT TGG GGC   1320
Arg Pro Arg Glu Ser Lys Ser Cys Gly Gly Pro Asp Ser Phe Tyr Cys Ala Tyr Trp Gly

TGT GAG ACA ACC GGT AGA GCT TAC TGG AAG CCC TCC TCA TCA TGG GAT TTC ATC ACA GTA   1380
Cys Glu Thr Thr Gly Arg Ala Tyr Trp Lys Pro Ser Ser Ser Trp Asp Phe Ile Thr Val

AAC AAC AAT CTC ACC TCT GAC CAG GCT GTC CAG GTA TGC AAA GAT AAT AAG TGG TGC AAC   1440
Asn Asn Asn Leu Thr Ser Asp Gln Ala Val Gln Val Cys Lys Asp Asn Lys Trp Cys Asn

CCC TTA GTT ATT CGG TTT ACA GAC GCC GGG AGA CGG GTT ACT TCC TGG ACC ACA GGA CAT   1500
Pro Leu Val Ile Arg Phe Thr Asp Ala Gly Arg Arg Val Thr Ser Trp Thr Thr Gly His

TAC TGG GGC TTA CGT TTG TAT GTC TCC GGA CAA GAT CCA GGG CTT ACA TTT GGG ATC AGA   1560
Tyr Trp Gly Leu Arg Leu Tyr Val Ser Gly Gln Asp Pro Gly Leu Thr Phe Gly Ile Arg

CTC AGA TAC CAA AAT CTA GGA CCC CGC GTC CCA ATA GGG ATC CAA TAC CCC TAC GAC GTC   1620
Leu Arg Tyr Gln Asn Leu Gly Pro Arg Val Pro Ile Gly Ile Gln Tyr Pro Tyr Asp Val

CCC GAC TAT GCC AAT CGA TTA GTT CAA TTT GTT AAA GAC AGG ATC TCA GTA GTC CAG GCT   1680
Pro Asp Tyr Ala Asn Arg Leu Val Gln Phe Val Lys Asp Arg Ile Ser Val Val Gln Ala

TTA GTC CTG ACT CAA CAA TAC CAC CAG CTA AAA CCA CTA GAA TAC GAG CCA TGA           1734
Leu Val Leu Thr Gln Gln Tyr His Gln Leu Lys Pro Leu Glu Tyr Glu Pro ...
```

FIG. 13 (continued)

34.1-GALV-RBD- Sequence

```
              10          20          30          40          50          60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
ATG GCG CGT TCA ACG CTC TCA AAA CCC CTT AAA AAT AAG GTT AAC CCG CGA GGC CCC CTA     60
Met Ala Arg Ser Thr Leu Ser Lys Pro Leu Lys Asn Lys Val Asn Pro Arg Gly Pro Leu

ATC CCC TTA ATT CTT CTG ATG CTC AGA GGG GTC AGT ACT GCT TCG CCC GGC TCC AGT GCG    120
Ile Pro Leu Ile Leu Leu Met Leu Arg Gly Val Ser Thr Ala Ser Pro Gly Ser Ser Ala

GCC CAG CCG GCC ATG GCC CAG GTT CAG CTT CAG CAG TCT GGA GCT GAA CTG GTA AAG CCT    180
Ala Gln Pro Ala Met Ala Gln Val Gln Leu Gln Gln Ser Gly Ala Glu Leu Val Lys Pro

GGG GCT TCA GTG AAG TTG TCC TGC AAG GCT TCT GGC TAC ACC TTC ACA AGC TAT GAT ATA    240
Gly Ala Ser Val Lys Leu Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr Asp Ile

AAC TGG GTG AGG CAG AGG CCT GAA CAG GGA CTT GAG TGG ATT GGA TGG ATT TTT CCT GGA    300
Asn Trp Val Arg Gln Arg Pro Glu Gln Gly Leu Glu Trp Ile Gly Trp Ile Phe Pro Gly

GAT GGT AGT ACT AAG TAC AAT GAG AAG TTC AAG GGC AAG GCC ACA CTG ACT ACA GAC AAA    360
Asp Gly Ser Thr Lys Tyr Asn Glu Lys Phe Lys Gly Lys Ala Thr Leu Thr Thr Asp Lys

TCC TCC AGC ACA GCC TAC ATG CAG CTC AGC AGG CTG ACA TCT GAA GAC TCT GCT ATC TAT    420
Ser Ser Ser Thr Ala Tyr Met Gln Leu Ser Arg Leu Thr Ser Glu Asp Ser Ala Ile Tyr

TTC TGT GCA AGA ATG GCC GAC TAT GGT CAC TAC GGC GCT GTG GAC TAC TGG GGT CAA GGG    480
Phe Cys Ala Arg Met Ala Asp Tyr Gly His Tyr Gly Ala Val Asp Tyr Trp Gly Gln Gly

ACC ACG GTC ACC GTC TCC TCA GGT GGA GGC GGT TCA GGC GGA GGT GGC TCT GGC GGT GGC    540
Thr Thr Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly

GGA TCG GAC ATC CAG ATG ACT CAG TCT CCA TCC TCC CTG TCT GCC TCT CTG GGA GAC AGA    600
Gly Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Leu Gly Asp Arg

GTC ACC ATC AGT TGC AGG GCA AGT CAG GAC ATT AGC AAT TAT TTA AAC TGG TAT CAG CAG    660
Val Thr Ile Ser Cys Arg Ala Ser Gln Asp Ile Ser Asn Tyr Leu Asn Trp Tyr Gln Gln

ACA CCA GAT GGA ACT GTT AAA CTC CTG ATC TAC TAC ACA TCA AGA TTA CAC TCA GGA GTC    720
Thr Pro Asp Gly Thr Val Lys Leu Leu Ile Tyr Tyr Thr Ser Arg Leu His Ser Gly Val

CCA TCA AGG TTC AGT GGC AGT GGG TCT GGA ACA GAT TAT TCT CTC ACC ATT AGC AAC CTG    780
Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Tyr Ser Leu Thr Ile Ser Asn Leu

GAG CAA GAA GAT ATT GCC ACT TAC TTT TGC CAA CAG GGT TAT ACG CTT CCG TGG ACG TTC    840
Glu Gln Glu Asp Ile Ala Thr Tyr Phe Cys Gln Gln Gly Tyr Thr Leu Pro Trp Thr Phe

GGT GGA GGC ACC AAG CTG GAG ATC AAA CGG GCG GCC GCA ATC GAG GGA AGG AGT CTG CAA    900
Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg Ala Ala Ala Ile Glu Gly Arg Ser Leu Gln

AAT AAG AAC CCC CAC CAG CCC ATG ACC CTC ACT TGG CAG GTA CTG TCC CAA ACT GGA GAC    960
Asn Lys Asn Pro His Gln Pro Met Thr Leu Thr Trp Gln Val Leu Ser Gln Thr Gly Asp

GTT GTC TGG GAT ACA AAG GCA GTC CAG CCC CCT TGG ACT TGG TGG CCC ACA CTT AAA CCT   1020
Val Val Trp Asp Thr Lys Ala Val Gln Pro Pro Trp Thr Trp Trp Pro Thr Leu Lys Pro

GAT GTA TGT GCC TTG GCG GCT AGT CTT GAG TCC TGG GAT ATC CCG GGA ACC GAT GTC TCG   1080
Asp Val Cys Ala Leu Ala Ala Ser Leu Glu Ser Trp Asp Ile Pro Gly Thr Asp Val Ser
```

FIG. 14

34.1-GALV-RBD- Sequence

```
                10                  20                  30                  40                  50                  60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
TCC TCT AAA CGA GTC AGA CCT CCG GAC TCA GAC TAT ACT GCC GCT TAT AAG CAA ATC ACC   1140
Ser Ser Lys Arg Val Arg Pro Pro Asp Ser Asp Tyr Thr Ala Ala Tyr Lys Gln Ile Thr

TGG GGA GCC ATA GGG TGC AGC TAC CCT CGG GCT AGG ACT AGA ATG GCA AGC TCT ACC TTC   1200
Trp Gly Ala Ile Gly Cys Ser Tyr Pro Arg Ala Arg Thr Arg Met Ala Ser Ser Thr Phe

TAC GTA TGT CCC CGG GAT GGC CGG ACC CTT TCA GAA GCT AGA AGG TGC GGG GGG CTA GAA   1260
Tyr Val Cys Pro Arg Asp Gly Arg Thr Leu Ser Glu Ala Arg Arg Cys Gly Gly Leu Glu

TCC CTA TAC TGT AAA GAA TGG GAT TGT GAG ACC ACG GGG ACC GGT TAT TGG CTA TCT AAA   1320
Ser Leu Tyr Cys Lys Glu Trp Asp Cys Glu Thr Thr Gly Thr Gly Tyr Trp Leu Ser Lys

TCC TCA AAA GAC CTC ATA ACT GTA AAA TGG GAC CAA AAT AGC GAA TGG ACT CAA AAA TTT   1380
Ser Ser Lys Asp Leu Ile Thr Val Lys Trp Asp Gln Asn Ser Glu Trp Thr Gln Lys Phe

CAA CAG TGT CAC CAG ACC GGC TGG TGT AAC CCC CTT AAA ATA GAT TTC ACA GAC AAA GGA   1440
Gln Gln Cys His Gln Thr Gly Trp Cys Asn Pro Leu Lys Ile Asp Phe Thr Asp Lys Gly

AAA TTA TCC AAG GAC TGG ATA ACG GGA AAA ACC TGG GGA TTA AGA TTC TAT GTG TCT GGA   1500
Lys Leu Ser Lys Asp Trp Ile Thr Gly Lys Thr Trp Gly Leu Arg Phe Tyr Val Ser Gly

CAT CCA GGC GTA CAG TTC ACC ATT CGC TTA AAA ATC ACC AAC ATG CCA GCT GTG GCA GTA   1560
His Pro Gly Val Gln Phe Thr Ile Arg Leu Lys Ile Thr Asn Met Pro Ala Val Ala Val

GGG ATC CAA TAC CCC TAC GAC GTC CCC GAC TAT GCC AAT CGA TTA GTT CAA TTT GTT AAA   1620
Gly Ile Gln Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Asn Arg Leu Val Gln Phe Val Lys

GAC AGG ATC TCA GTA GTC CAG GCT TTA GTC CTG ACT CAA CAA TAC CAC CAG CTA AAA CCA   1680
Asp Arg Ile Ser Val Val Gln Ala Leu Val Leu Thr Gln Gln Tyr His Gln Leu Lys Pro

CTA GAA TAC GAG CCA TGA                                                           1698
Leu Glu Tyr Glu Pro ...
```

FIG. 14 (continued)

AdelH Sequence

```
                10          20          30          40          50          60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
ATG GCG CGT TCA ACG CTC TCA AAA CCC CCT CAA GAT AAG ATT AAC CCG TGG AAG CCC TTA     60
Met Ala Arg Ser Thr Leu Ser Lys Pro Pro Gln Asp Lys Ile Asn Pro Trp Lys Pro Leu

ATA GTC ATG GGA GTC CTG TTA GGA GTA GGG ATG GCA GAG AGC CCC CAG GTC TTT AAT GTA    120
Ile Val Met Gly Val Leu Leu Gly Val Gly Met Ala Glu Ser Pro Gln Val Phe Asn Val

ACC TGG AGA GTC ACC AAC CTG ATG ACT GGG CGT ACC GCC AAT GCC ACC TCC CTC CTG GGA    180
Thr Trp Arg Val Thr Asn Leu Met Thr Gly Arg Thr Ala Asn Ala Thr Ser Leu Leu Gly

ACT GTA CAA GAT GCC TTC CCA AAA TTA TAT TTT GAT CTA TGT GAT CTG GTC GGA GAG GAG    240
Thr Val Gln Asp Ala Phe Pro Lys Leu Tyr Phe Asp Leu Cys Asp Leu Val Gly Glu Glu

TGG GAC CCT TCA GAC CAG GAA CCG TAT GTC GGG TAT GGC TGC AAG TAC CCC GCA GGG AGA    300
Trp Asp Pro Ser Asp Gln Glu Pro Tyr Val Gly Tyr Gly Cys Lys Tyr Pro Ala Gly Arg

CAG CGG ACC CGG ACT TTT GAC TTT TAC GTG TGC CCT GGG CAT ACC GTA AAG TCG GGG TGT    360
Gln Arg Thr Arg Thr Phe Asp Phe Tyr Val Cys Pro Gly His Thr Val Lys Ser Gly Cys

GGG GGA CCA GGA GAG GGC TAC TGT GGT AAA TGG GGG TGT GAA ACC ACC GGA CAG GCT TAC    420
Gly Gly Pro Gly Glu Gly Tyr Cys Gly Lys Trp Gly Cys Glu Thr Thr Gly Gln Ala Tyr

TGG AAG CCC ACA TCA TCG TGG GAC CTA ATC TCC CTT AAG CGC GGT AAC ACC CCC TGG GAC    480
Trp Lys Pro Thr Ser Ser Trp Asp Leu Ile Ser Leu Lys Arg Gly Asn Thr Pro Trp Asp

ACG GGA TGC TCT AAA GTT GCC TGT GGC CCC TGC TAC GAC CTC TCC AAA GTA TCC AAT TCC    540
Thr Gly Cys Ser Lys Val Ala Cys Gly Pro Cys Tyr Asp Leu Ser Lys Val Ser Asn Ser

TTC CAA GGG GCT ACT CGA GGG GGC AGA TGC AAC CCT CTA GTC CTA GAA TTC ACT GAT GCA    600
Phe Gln Gly Ala Thr Arg Gly Gly Arg Cys Asn Pro Leu Val Leu Glu Phe Thr Asp Ala

GGA AAA AAG GCT AAC TGG GAC GGG CCC AAA TCG TGG GGA CTG AGA CTG TAC CGG ACA GGA    660
Gly Lys Lys Ala Asn Trp Asp Gly Pro Lys Ser Trp Gly Leu Arg Leu Tyr Arg Thr Gly

ACA GAT CCT ATT ACC ATG TTC TCC CTG ACC CGG CAG GTC CTT AAT GTG GGA CCC CGA GTC    720
Thr Asp Pro Ile Thr Met Phe Ser Leu Thr Arg Gln Val Leu Asn Val Gly Pro Arg Val

CCC ATA GGG CCC AAC CCA GTA TTA CCC GAC CAA AGA CTC CCT TCC TCA CCA ATA GAG ATT    780
Pro Ile Gly Pro Asn Pro Val Leu Pro Asp Gln Arg Leu Pro Ser Ser Pro Ile Glu Ile

GTA CCG GCT CCA CAG CCA CCT AGC CCC CTC AAT ACC AGT TAC CCC CCT TCC ACT ACC AGT    840
Val Pro Ala Pro Gln Pro Pro Ser Pro Leu Asn Thr Ser Tyr Pro Pro Ser Thr Thr Ser

ACA CCC TCA ACC TCC CCT ACA AGT CCA AGT GTC CCA CAG CCA CCC CCA GGA ACT GGA GAT    900
Thr Pro Ser Thr Ser Pro Thr Ser Pro Ser Val Pro Gln Pro Pro Pro Gly Thr Gly Asp

AGA CTA CTA GCT CTA GTC AAA GGA GCC TAT CAG GCG CTT AAC CTC ACC AAT CCC GAC AAG    960
Arg Leu Leu Ala Leu Val Lys Gly Ala Tyr Gln Ala Leu Asn Leu Thr Asn Pro Asp Lys

ACC CAA GAA TGT TGG CTG TGC TTA GTG TCG GGA CCT CCT TAT TAC GAA GGA GTA GCG GTC   1020
Thr Gln Glu Cys Trp Leu Cys Leu Val Ser Gly Pro Pro Tyr Tyr Glu Gly Val Ala Val

GTG GGC ACT TAT ACC AAT CAT TCC ACC GCT CCG GCC AAC TGT ACG GCC ACT TCC CAA CAT   1080
Val Gly Thr Tyr Thr Asn His Ser Thr Ala Pro Ala Asn Cys Thr Ala Thr Ser Gln His

AAG CTT ACC CTA TCT GAA GTG ACA GGA CAG GGC CTA TGC ATG GGG GCA GTA CCT AAA ACT   1140
Lys Leu Thr Leu Ser Glu Val Thr Gly Gln Gly Leu Cys Met Gly Ala Val Pro Lys Thr

CAC CAG GCC TTA TGT AAC ACC ACC CAA AGC GCC GGC TCA GGA TCC TAC TAC CTT GCA GCA   1200
His Gln Ala Leu Cys Asn Thr Thr Gln Ser Ala Gly Ser Gly Ser Tyr Tyr Leu Ala Ala
```

FIG. 15

AdelH Sequence

```
              10          20          30          40          50          60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
CCC GCC GGA ACA ATG TGG GCT TGC AGC ACT GGA TTG ACT CCC TGC TTG TCC ACC ACG GTG   1260
Pro Ala Gly Thr Met Trp Ala Cys Ser Thr Gly Leu Thr Pro Cys Leu Ser Thr Thr Val

CTC AAT CTA ACC ACA GAT TAT TGT GTA TTA GTT GAA CTC TGG CCC AGA GTA ATT TAC CAC   1320
Leu Asn Leu Thr Thr Asp Tyr Cys Val Leu Val Glu Leu Trp Pro Arg Val Ile Tyr His

TCC CCC GAT TAT ATG TAT GGT CAG CTT GAA CAG CGT ACC AAA TAT AAA AGA GAG CCA GTA   1380
Ser Pro Asp Tyr Met Tyr Gly Gln Leu Glu Gln Arg Thr Lys Tyr Lys Arg Glu Pro Val

TCA TTG ACC CTG GCC CTT CTA CTA GGA GGA TTA ACC ATG GGA GGG ATT GCA GCT GGA ATA   1440
Ser Leu Thr Leu Ala Leu Leu Leu Gly Gly Leu Thr Met Gly Gly Ile Ala Ala Gly Ile

GGG ACG GGG ACC ACT GCC TTA ATT AAA ACC CAG CAG TTT GAG CAG CTT CAT GCC GCT ATC   1500
Gly Thr Gly Thr Thr Ala Leu Ile Lys Thr Gln Gln Phe Glu Gln Leu His Ala Ala Ile

CAG ACA GAC CTC AAC GAA GTC GAA AAG TCA ATT ACC AAC CTA GAA AAG TCA CTG ACC TCG   1560
Gln Thr Asp Leu Asn Glu Val Glu Lys Ser Ile Thr Asn Leu Glu Lys Ser Leu Thr Ser

TTG TCT GAA GTA GTC CTA CAG AAC CGC AGA GGC CTA GAT TTG CTA TTC CTA AAG GAG GGA   1620
Leu Ser Glu Val Val Leu Gln Asn Arg Arg Gly Leu Asp Leu Leu Phe Leu Lys Glu Gly

GGT CTC TGC GCA GCC CTA AAA GAA GAA TGT TGT TTT TAT GCA GAC CAC ACG GGG CTA GTG   1680
Gly Leu Cys Ala Ala Leu Lys Glu Glu Cys Cys Phe Tyr Ala Asp His Thr Gly Leu Val

AGA GAC AGC ATG GCC AAA TTA AGA GAA AGG CTT AAT CAG AGA CAA AAA CTA TTT GAG ACA   1740
Arg Asp Ser Met Ala Lys Leu Arg Glu Arg Leu Asn Gln Arg Gln Lys Leu Phe Glu Thr

GGC CAA GGA TGG TTC GAA GGG CTG TTT AAT AGA TCC CCC TGG TTT ACC ACC TTA ATC TCC   1800
Gly Gln Gly Trp Phe Glu Gly Leu Phe Asn Arg Ser Pro Trp Phe Thr Thr Leu Ile Ser

ACC ATC ATG GGA CCT CTA ATA GTA CTC TTA CTG ATC TTA CTC TTT GGA CCT TGC ATT CTC   1860
Thr Ile Met Gly Pro Leu Ile Val Leu Leu Leu Ile Leu Leu Phe Gly Pro Cys Ile Leu

AAT CGA TTA GTT CAA TTT GTT AAA GAC AGG ATC TCA GTA GTC CAG GCT TTA GTC CTG ACT   1920
Asn Arg Leu Val Gln Phe Val Lys Asp Arg Ile Ser Val Val Gln Ala Leu Val Leu Thr

CAA CAA TAC CAC CAG CTA AAA CCA CTA GAA TAC GAG CCA                               1959
Gln Gln Tyr His Gln Leu Lys Pro Leu Glu Tyr Glu Pro
```

FIG. 15 (continued)

PROMOdelH Sequence

```
              10                  20                  30                  40                  50                  60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
ATG GCG CGT TCA ACG CTC TCA AAA CCC CCT CAA GAT AAG ATT AAC CCG TGG AAG CCC TTA    60
Met Ala Arg Ser Thr Leu Ser Lys Pro Pro Gln Asp Lys Ile Asn Pro Trp Lys Pro Leu

ATA GTC ATG GGA GTC CTG TTA GGA GTA GGG ATG GCA GAG AGC CCC CAG GTC TTT AAT GTA   120
Ile Val Met Gly Val Leu Leu Gly Val Gly Met Ala Glu Ser Pro Gln Val Phe Asn Val

ACC TGG AGA GTC ACC AAC CTG ATG ACT GGG CGT ACC GCC AAT GCC ACC TCC CTC CTG GGA   180
Thr Trp Arg Val Thr Asn Leu Met Thr Gly Arg Thr Ala Asn Ala Thr Ser Leu Leu Gly

ACT GTA CAA GAT GCC TTC CCA AAA TTA TAT TTT GAT CTA TGT GAT CTG GTC GGA GAG GAG   240
Thr Val Gln Asp Ala Phe Pro Lys Leu Tyr Phe Asp Leu Cys Asp Leu Val Gly Glu Glu

TGG GAC CCT TCA GAC CAG GAA CCG TAT GTC GGG TAT GGC TGC AAG TAC CCC GCA GGG AGA   300
Trp Asp Pro Ser Asp Gln Glu Pro Tyr Val Gly Tyr Gly Cys Lys Tyr Pro Ala Gly Arg

CAG CGG ACC CGG ACT TTT GAC TTT TAC GTG TGC CCT GGG CAT ACC GTA AAG TCG GGG TGT   360
Gln Arg Thr Arg Thr Phe Asp Phe Tyr Val Cys Pro Gly His Thr Val Lys Ser Gly Cys

GGG GGA CCA GGA GAG GGC TAC TGT GGT AAA TGG GGG TGT GAA ACC ACC GGA CAG GCT TAC   420
Gly Gly Pro Gly Glu Gly Tyr Cys Gly Lys Trp Gly Cys Glu Thr Thr Gly Gln Ala Tyr

TGG AAG CCC ACA TCA TCG TGG GAC CTA ATC TCC CTT AAG CGC GGT AAC ACC CCC TGG GAC   480
Trp Lys Pro Thr Ser Ser Trp Asp Leu Ile Ser Leu Lys Arg Gly Asn Thr Pro Trp Asp

ACG GGA TGC TCT AAA GTT GCC TGT GGC CCC TGC TAC GAC CTC TCC AAA GTA TCC AAT TCC   540
Thr Gly Cys Ser Lys Val Ala Cys Gly Pro Cys Tyr Asp Leu Ser Lys Val Ser Asn Ser

TTC CAA GGG GCT ACT CGA GGG GGC AGA TGC AAC CCT CTA GTC CTA GAA TTC ACT GAT GCA   600
Phe Gln Gly Ala Thr Arg Gly Gly Arg Cys Asn Pro Leu Val Leu Glu Phe Thr Asp Ala

GGA AAA AAG GCT AAC TGG GAC GGG CCC AAA TCG TGG GGA CTG AGA CTG TAC CGG ACA GGA   660
Gly Lys Lys Ala Asn Trp Asp Gly Pro Lys Ser Trp Gly Leu Arg Leu Tyr Arg Thr Gly

ACA GAT CCT ATT ACC ATG TTC TCC CTG ACC CGG CAG GTC CTT AAT GTG GGA CCC CGA GTC   720
Thr Asp Pro Ile Thr Met Phe Ser Leu Thr Arg Gln Val Leu Asn Val Gly Pro Arg Val

CCC ATA GGG CCC AAC CCC GTT CTG GCA GAC CAA CAG CCA CTC TCC AAG CCC AAA CCT GTT   780
Pro Ile Gly Pro Asn Pro Val Leu Ala Asp Gln Gln Pro Leu Ser Lys Pro Lys Pro Val

AAG TCG CCT TCA GTC ACC AAA CCA CCC AGT GGG ACT CCT CTC TCC CCT ACC CAA CTT CCA   840
Lys Ser Pro Ser Val Thr Lys Pro Pro Ser Gly Thr Pro Leu Ser Pro Thr Gln Leu Pro

CCG GCG GGT ACC GGA GAT AGA CTA CTA GCT CTA GTC AAA GGA GCC TAT CAG GCG CTT AAC   900
Pro Ala Gly Thr Gly Asp Arg Leu Leu Ala Leu Val Lys Gly Ala Tyr Gln Ala Leu Asn

CTC ACC AAT CCC GAC AAG ACC CAA GAA TGT TGG CTG TGC TTA GTG TCG GGA CCT CCT TAT   960
Leu Thr Asn Pro Asp Lys Thr Gln Glu Cys Trp Leu Cys Leu Val Ser Gly Pro Pro Tyr

TAC GAA GGA GTA GCG GTC GTG GGC ACT TAT ACC AAT CAT TCC ACC GCT CCG GCC AAC TGT  1020
Tyr Glu Gly Val Ala Val Val Gly Thr Tyr Thr Asn His Ser Thr Ala Pro Ala Asn Cys

ACG GCC ACT TCC CAA CAT AAG CTT ACC CTA TCT GAA GTG ACA GGA CAG GGC CTA TGC ATG  1080
Thr Ala Thr Ser Gln His Lys Leu Thr Leu Ser Glu Val Thr Gly Gln Gly Leu Cys Met

GGG GCA GTA CCT AAA ACT CAC CAG GCC TTA TGT AAC ACC ACC CAA AGC GCC GGC TCA GGA  1140
Gly Ala Val Pro Lys Thr His Gln Ala Leu Cys Asn Thr Thr Gln Ser Ala Gly Ser Gly

TCC TAC TAC CTT GCA GCA CCC GCC GGA ACA ATG TGG GCT TGC AGC ACT GGA TTG ACT CCC  1200
Ser Tyr Tyr Leu Ala Ala Pro Ala Gly Thr Met Trp Ala Cys Ser Thr Gly Leu Thr Pro
```

FIG. 16

PROMOdelH Sequence

```
                10                  20                  30                  40                  50                  60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
TGC TTG TCC ACC ACG GTG CTC AAT CTA ACC ACA GAT TAT TGT GTA TTA GTT GAA CTC TGG   1260
Cys Leu Ser Thr Thr Val Leu Asn Leu Thr Thr Asp Tyr Cys Val Leu Val Glu Leu Trp

CCC AGA GTA ATT TAC CAC TCC CCC GAT TAT ATG TAT GGT CAG CTT GAA CAG CGT ACC AAA   1320
Pro Arg Val Ile Tyr His Ser Pro Asp Tyr Met Tyr Gly Gln Leu Glu Gln Arg Thr Lys

TAT AAA AGA GAG CCA GTA TCA TTG ACC CTG GCC CTT CTA CTA GGA GGA TTA ACC ATG GGA   1380
Tyr Lys Arg Glu Pro Val Ser Leu Thr Leu Ala Leu Leu Leu Gly Gly Leu Thr Met Gly

GGG ATT GCA GCT GGA ATA GGG ACG GGG ACC ACT GCC TTA ATT AAA ACC CAG CAG TTT GAG   1440
Gly Ile Ala Ala Gly Ile Gly Thr Gly Thr Thr Ala Leu Ile Lys Thr Gln Gln Phe Glu

CAG CTT CAT GCC GCT ATC CAG ACA GAC CTC AAC GAA GTC GAA AAG TCA ATT ACC AAC CTA   1500
Gln Leu His Ala Ala Ile Gln Thr Asp Leu Asn Glu Val Glu Lys Ser Ile Thr Asn Leu

GAA AAG TCA CTG ACC TCG TTG TCT GAA GTA GTC CTA CAG AAC CGC AGA GGC CTA GAT TTG   1560
Glu Lys Ser Leu Thr Ser Leu Ser Glu Val Val Leu Gln Asn Arg Arg Gly Leu Asp Leu

CTA TTC CTA AAG GAG GGA GGT CTC TGC GCA GCC CTA AAA GAA GAA TGT TGT TTT TAT GCA   1620
Leu Phe Leu Lys Glu Gly Gly Leu Cys Ala Ala Leu Lys Glu Glu Cys Cys Phe Tyr Ala

GAC CAC ACG GGG CTA GTG AGA GAC AGC ATG GCC AAA TTA AGA GAA AGG CTT AAT CAG AGA   1680
Asp His Thr Gly Leu Val Arg Asp Ser Met Ala Lys Leu Arg Glu Arg Leu Asn Gln Arg

CAA AAA CTA TTT GAG ACA GGC CAA GGA TGG TTC GAA GGG CTG TTT AAT AGA TCC CCC TGG   1740
Gln Lys Leu Phe Glu Thr Gly Gln Gly Trp Phe Glu Gly Leu Phe Asn Arg Ser Pro Trp

TTT ACC ACC TTA ATC TCC ACC ATC ATG GGA CCT CTA ATA GTA CTC TTA CTG ATC TTA CTC   1800
Phe Thr Thr Leu Ile Ser Thr Ile Met Gly Pro Leu Ile Val Leu Leu Leu Ile Leu Leu

TTT GGA CCT TGC ATT CTC AAT CGA TTA GTT CAA TTT GTT AAA GAC AGG ATC TCA GTA GTC   1860
Phe Gly Pro Cys Ile Leu Asn Arg Leu Val Gln Phe Val Lys Asp Arg Ile Ser Val Val

CAG GCT TTA GTC CTG ACT CAA CAA TAC CAC CAG CTA AAA CCA CTA GAA TAC GAG CCA       1917
Gln Ala Leu Val Leu Thr Gln Gln Tyr His Gln Leu Lys Pro Leu Glu Tyr Glu Pro
```

FIG. 16 (continued)

GALVdelH Sequence

```
                 10                  20                  30                  40                  50                  60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
ATG GTA TTG CTG CCT GGG TCC ATG CTT CTC ACC TCA AAC CTG CAC CAC CTT CGG CAC CAG     60
Met Val Leu Leu Pro Gly Ser Met Leu Leu Thr Ser Asn Leu His His Leu Arg His Gln

ATG AGT CCT GGG AGC TGG AAA AGA CTG ATC ATC CTC TTA AGC TGC GTA TTC GGC GGC GGC    120
Met Ser Pro Gly Ser Trp Lys Arg Leu Ile Ile Leu Leu Ser Cys Val Phe Gly Gly Gly

GGG ACG AGT CTG CAG AAT AAG AAC CCC CAG CCC ATG ACC CTC ACT TGG CAG GTA CTG TCC    180
Gly Thr Ser Leu Gln Asn Lys Asn Pro Gln Pro Met Thr Leu Thr Trp Gln Val Leu Ser

CAA ACT GGA GAC GTT GTC TGG GAT ACA AAG GCA GTC CAG CCC CCT TGG ACT TGG TGG CCC    240
Gln Thr Gly Asp Val Val Trp Asp Thr Lys Ala Val Gln Pro Pro Trp Thr Trp Trp Pro

ACA CTT AAA CCT GAT GTA TGT GCC TTG GCG GCT AGT CTT GAG TCC TGG GAT ATC CCG GGA    300
Thr Leu Lys Pro Asp Val Cys Ala Leu Ala Ala Ser Leu Glu Ser Trp Asp Ile Pro Gly

ACC GAT GTC TCG TCC TCT AAA CGA GTC AGA CCT CCG GAC TCA GAC TAT ACT GCC GCT TAT    360
Thr Asp Val Ser Ser Ser Lys Arg Val Arg Pro Pro Asp Ser Asp Tyr Thr Ala Ala Tyr

AAG CAA ATC ACC TGG GGA GCC ATA GGG TGC AGC TAC CCT CGG GCT AGG ACT AGA ATG GCA    420
Lys Gln Ile Thr Trp Gly Ala Ile Gly Cys Ser Tyr Pro Arg Ala Arg Thr Arg Met Ala

AGC TCT ACC TTC TAC GTA TGT CCC CGG GAT GGC CGG ACC CTT TCA GAA GCT AGA AGG TGC    480
Ser Ser Thr Phe Tyr Val Cys Pro Arg Asp Gly Arg Thr Leu Ser Glu Ala Arg Arg Cys

GGG GGG CTA GAA TCC CTA TAC TGT AAA GAA TGG GAT TGT GAG ACC ACG GGG ACC GGT TAT    540
Gly Gly Leu Glu Ser Leu Tyr Cys Lys Glu Trp Asp Cys Glu Thr Thr Gly Thr Gly Tyr

TGG CTA TCT AAA TCC TCA AAA GAC CTC ATA ACT GTA AAA TGG GAC CAA AAT AGC GAA TGG    600
Trp Leu Ser Lys Ser Ser Lys Asp Leu Ile Thr Val Lys Trp Asp Gln Asn Ser Glu Trp

ACT CAA AAA TTT CAA CAG TGT CAC CAG ACC GGC TGG TGT AAC CCC CTT AAA ATA GAT TTC    660
Thr Gln Lys Phe Gln Gln Cys His Gln Thr Gly Trp Cys Asn Pro Leu Lys Ile Asp Phe

ACA GAC AAA GGA AAA TTA TCC AAG GAC TGG ATA ACG GGA AAA ACC TGG GGA TTA AGA TTC    720
Thr Asp Lys Gly Lys Leu Ser Lys Asp Trp Ile Thr Gly Lys Thr Trp Gly Leu Arg Phe

TAT GTG TCT GGA CAT CCA GGC GTA CAG TTC ACC ATT CGC TTA AAA ATC ACC AAC ATG CCA    780
Tyr Val Ser Gly His Pro Gly Val Gln Phe Thr Ile Arg Leu Lys Ile Thr Asn Met Pro

GCT GTG GCA GTA GGT CCT GAC CTC GTC CTT GTG GAA CAA GGA CCT CCT AGA ACG TCC CTC    840
Ala Val Ala Val Gly Pro Asp Leu Val Leu Val Glu Gln Gly Pro Pro Arg Thr Ser Leu

GCT CTC CCA CCT CCT CTT CCC CCA AGG GAA GCG CCA CCG CCA TCT CTC CCC GAC TCT AAC    900
Ala Leu Pro Pro Pro Leu Pro Pro Arg Glu Ala Pro Pro Pro Ser Leu Pro Asp Ser Asn

TCC ACA GCC CTG GCG ACT AGT GCA CAA ACT CCC ACG GTG AGA AAA ACA ATT GTT ACC CTA    960
Ser Thr Ala Leu Ala Thr Ser Ala Gln Thr Pro Thr Val Arg Lys Thr Ile Val Thr Leu

AAC ACT CCG CCT CCC ACC ACA GGC GAC AGA CTT TTT GAT CTT GTG CAG GGG GCC TTC CTA   1020
Asn Thr Pro Pro Pro Thr Thr Gly Asp Arg Leu Phe Asp Leu Val Gln Gly Ala Phe Leu

ACC TTA AAT GCT ACC AAC CCA GGG GCC ACT GAG TCT TGC TGG CTT TGT TTG GCC ATG GGC   1080
Thr Leu Asn Ala Thr Asn Pro Gly Ala Thr Glu Ser Cys Trp Leu Cys Leu Ala Met Gly

CCC CCT TAT TAT GAA GCA ATA GCC TCA TCA GGA GAG GTC GCC TAC TCC ACC GAC CTT GAC   1140
Pro Pro Tyr Tyr Glu Ala Ile Ala Ser Ser Gly Glu Val Ala Tyr Ser Thr Asp Leu Asp

CGG TGC CGC TGG GGG ACC CAA GGA AAG CTC ACC CTC ACT GAG GTC TCA GGA CAC GGG TTG   1200
Arg Cys Arg Trp Gly Thr Gln Gly Lys Leu Thr Leu Thr Glu Val Ser Gly His Gly Leu
```

FIG. 17

GALVdelH Sequence

```
                10          20          30          40          50          60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
TGC ATA GGA AAG GTG CCC TTT ACC CAT CAG CAT CTC TGC AAT CAG ACC CTA TCC ATC AAT   1260
Cys Ile Gly Lys Val Pro Phe Thr His Gln His Leu Cys Asn Gln Thr Leu Ser Ile Asn

TCC TCC GGA GAC CAT CAG TAT CTG CTC CCC TCC AAC CAT AGC TGG TGG GCT TGC AGC ACT   1320
Ser Ser Gly Asp His Gln Tyr Leu Leu Pro Ser Asn His Ser Trp Trp Ala Cys Ser Thr

GGC CTC ACC CCT TGC CTC TCC ACC TCA GTT TTT AAT CAG ACT AGA GAT TTC TGT ATC CAG   1380
Gly Leu Thr Pro Cys Leu Ser Thr Ser Val Phe Asn Gln Thr Arg Asp Phe Cys Ile Gln

GTC CAG CTG ATT CCT CGC ATC TAT TAC TAT CCT GAA GAA GTT TTG TTA CAG GCC TAT GAC   1440
Val Gln Leu Ile Pro Arg Ile Tyr Tyr Tyr Pro Glu Glu Val Leu Leu Gln Ala Tyr Asp

AAT TCT CAC CCC AGG ACT AAA AGA GAG GCT GTC TCA CTT ACC CTA GCT GTT TTA CTG GGG   1500
Asn Ser His Pro Arg Thr Lys Arg Glu Ala Val Ser Leu Thr Leu Ala Val Leu Leu Gly

TTG GGA ATC ACG GCG GGA ATA GGT ACT GGT TCA ACT GCC TTA ATT AAA GGA CCT ATA GAC   1560
Leu Gly Ile Thr Ala Gly Ile Gly Thr Gly Ser Thr Ala Leu Ile Lys Gly Pro Ile Asp

CTC CAG CAA GGC CTG ACA AGC CTC CAG ATC GCC ATA GAT GCT GAC CTC CGG GCC CTC CAA   1620
Leu Gln Gln Gly Leu Thr Ser Leu Gln Ile Ala Ile Asp Ala Asp Leu Arg Ala Leu Gln

GAC TCA GTC AGC AAG TTA GAG GAC TCA CTG ACT TCC CTG TCC GAG GTA GTG CTC CAA AAT   1680
Asp Ser Val Ser Lys Leu Glu Asp Ser Leu Thr Ser Leu Ser Glu Val Val Leu Gln Asn

AGG AGA GGC CTT GAC TTG CTG TTT CTA AAA GAA GGT GGC CTC TGT GCG GCC CTA AAG GAA   1740
Arg Arg Gly Leu Asp Leu Leu Phe Leu Lys Glu Gly Gly Leu Cys Ala Ala Leu Lys Glu

GAG TGC TGT TTT TAC ATA GAC CAC TCA GGT GCA GTA CGG GAC TCC ATG AAA AAA CTC AAA   1800
Glu Cys Cys Phe Tyr Ile Asp His Ser Gly Ala Val Arg Asp Ser Met Lys Lys Leu Lys

GAA AAA CTG GAT AAA AGA CAG TTA GAG CGC CAG AAA AGC CAA AAC TGG TAT GAA GGA TGG   1860
Glu Lys Leu Asp Lys Arg Gln Leu Glu Arg Gln Lys Ser Gln Asn Trp Tyr Glu Gly Trp

TTC AAT AAC TCC CCT TGG TTC ACT ACC CTG CTA TCA ACC ATC GCT GGG CCC CTA TTA CTC   1920
Phe Asn Asn Ser Pro Trp Phe Thr Thr Leu Leu Ser Thr Ile Ala Gly Pro Leu Leu Leu

CTC CTT CTG TTG CTC ATC CTC GGG CCA TGC ATC ATC AAT CGA TTA GTT CAA TTT GTT AAA   1980
Leu Leu Leu Leu Leu Ile Leu Gly Pro Cys Ile Ile Asn Arg Leu Val Gln Phe Val Lys

GAC AGG ATC TCA GTA GTC CAG GCT TTA GTC CTG ACT CAA CAA TAC CAC CAG CTA AAG CCT   2040
Asp Arg Ile Ser Val Val Gln Ala Leu Val Leu Thr Gln Gln Tyr His Gln Leu Lys Pro

ATA GAG TAC GAG CCA                                                               2055
Ile Glu Tyr Glu Pro
```

FIG. 17 (continued)

C2delH Sequence

```
                     10                  20                  30                  40                  50                  60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
ATG GCG CGT TCA ACG CTC TCA AAA CCC CCT CAA GAT AAG ATT AAC CCG TGG AAG CCC TTA     60
Met Ala Arg Ser Thr Leu Ser Lys Pro Pro Gln Asp Lys Ile Asn Pro Trp Lys Pro Leu

ATA GTC ATG GGA GTC CTG TTA GGA GTA GGG ATG GCA GAG AGC CCC CAG GTC TTT AAT GTA    120
Ile Val Met Gly Val Leu Leu Gly Val Gly Met Ala Glu Ser Pro Gln Val Phe Asn Val

ACC TGG AGA GTC ACC AAC CTG ATG ACT GGG CGT ACC GCC AAT GCC ACC TCC CTC CTG GGA    180
Thr Trp Arg Val Thr Asn Leu Met Thr Gly Arg Thr Ala Asn Ala Thr Ser Leu Leu Gly

ACT GTA CAA GAT GCC TTC CCA AAA TTA TAT TTT GAT CTA TGT GAT CTG GTC GGA GAG GAG    240
Thr Val Gln Asp Ala Phe Pro Lys Leu Tyr Phe Asp Leu Cys Asp Leu Val Gly Glu Glu

TGG GAC CCT TCA GAC CAG GAA CCG TAT GTC GGG TAT GGC TGC AAG TAC CCC GCA GGG AGA    300
Trp Asp Pro Ser Asp Gln Glu Pro Tyr Val Gly Tyr Gly Cys Lys Tyr Pro Ala Gly Arg

CAG CGG ACC CGG ACT TTT GAC TTT TAC GTG TGC CCT GGG CAT ACC GTA AAG TCG GGG TGT    360
Gln Arg Thr Arg Thr Phe Asp Phe Tyr Val Cys Pro Gly His Thr Val Lys Ser Gly Cys

GGG GGA CCA GGA GAG GGC TAC TGT GGT AAA TGG GGG TGT GAA ACC ACC GGA CAG GCT TAC    420
Gly Gly Pro Gly Glu Gly Tyr Cys Gly Lys Trp Gly Cys Glu Thr Thr Gly Gln Ala Tyr

TGG AAG CCC ACA TCA TCG TGG GAC CTA ATC TCC CTT AAG CGC GGT AAC ACC CCC TGG GAC    480
Trp Lys Pro Thr Ser Ser Trp Asp Leu Ile Ser Leu Lys Arg Gly Asn Thr Pro Trp Asp

ACG GGA TGC TCT AAA GTT GCC TGT GGC CCC TGC TAC GAC CTC TCC AAA GTA TCC AAT TCC    540
Thr Gly Cys Ser Lys Val Ala Cys Gly Pro Cys Tyr Asp Leu Ser Lys Val Ser Asn Ser

TTC CAA GGG GCT ACT CGA GGG GGC AGA TGC AAC CCT CTA GTC CTA GAA TTC ACT GAT GCA    600
Phe Gln Gly Ala Thr Arg Gly Gly Arg Cys Asn Pro Leu Val Leu Glu Phe Thr Asp Ala

GGA AAA AAG GCT AAC TGG GAC GGG CCC AAA TCG TGG GGA CTG AGA CTG TAC CGG ACA GGA    660
Gly Lys Lys Ala Asn Trp Asp Gly Pro Lys Ser Trp Gly Leu Arg Leu Tyr Arg Thr Gly

ACA GAT CCT ATT ACC ATG TTC TCC CTG ACC CGG CAG GTC CTT AAT GTG GGA CCC CGA GTC    720
Thr Asp Pro Ile Thr Met Phe Ser Leu Thr Arg Gln Val Leu Asn Val Gly Pro Arg Val

CCC ATA GGG CCC AAC CCA GTA TTA CCC GAC CAA AGA CTC CCT TCC TCA CCA ATA GAG ATT    780
Pro Ile Gly Pro Asn Pro Val Leu Pro Asp Gln Arg Leu Pro Ser Ser Pro Ile Glu Ile

GTA CCG GCT CCA CAG CCA CCT AGC CCC CTC AAT ACC AGT TAC CCC CCT TCC ACT ACC AGT    840
Val Pro Ala Pro Gln Pro Pro Ser Pro Leu Asn Thr Ser Tyr Pro Pro Ser Thr Thr Ser

ACA CCC CCC CCA GGT ACC GGA GAT AGA CTA CTA GCT CTA GTC AAA GGA GCC TAT CAG GCG    900
Thr Pro Pro Pro Gly Thr Gly Asp Arg Leu Leu Ala Leu Val Lys Gly Ala Tyr Gln Ala

CTT AAC CTC ACC AAT CCC GAC AAG ACC CAA GAA TGT TGG CTG TGC TTA GTG TCG GGA CCT    960
Leu Asn Leu Thr Asn Pro Asp Lys Thr Gln Glu Cys Trp Leu Cys Leu Val Ser Gly Pro

CCT TAT TAC GAA GGA GTA GCG GTC GTG GGC ACT TAT ACC AAT CAT TCC ACC GCT CCG GCC   1020
Pro Tyr Tyr Glu Gly Val Ala Val Val Gly Thr Tyr Thr Asn His Ser Thr Ala Pro Ala

AAC TGT ACG GCC ACT TCC CAA CAT AAG CTT ACC CTA TCT GAA GTG ACA GGA CAG GGC CTA   1080
Asn Cys Thr Ala Thr Ser Gln His Lys Leu Thr Leu Ser Glu Val Thr Gly Gln Gly Leu

TGC ATG GGG GCA GTA CCT AAA ACT CAC CAG GCC TTA TGT AAC ACC ACC CAA AGC GCC GGC   1140
Cys Met Gly Ala Val Pro Lys Thr His Gln Ala Leu Cys Asn Thr Thr Gln Ser Ala Gly

TCA GGA TCC TAC TAC CTT GCA GCA CCC GCC GGA ACA ATG TGG GCT TGC AGC ACT GGA TTG   1200
Ser Gly Ser Tyr Tyr Leu Ala Ala Pro Ala Gly Thr Met Trp Ala Cys Ser Thr Gly Leu
```

FIG. 18

C2delH Sequence

```
                    10                  20                  30                  40                  50                  60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
ACT CCC TGC TTG TCC ACC ACG GTG CTC AAT CTA ACC ACA GAT TAT TGT GTA TTA GTT GAA   1260
Thr Pro Cys Leu Ser Thr Thr Val Leu Asn Leu Thr Thr Asp Tyr Cys Val Leu Val Glu

CTC TGG CCC AGA GTA ATT TAC CAC TCC CCC GAT TAT ATG TAT GGT CAG CTT GAA CAG CGT   1320
Leu Trp Pro Arg Val Ile Tyr His Ser Pro Asp Tyr Met Tyr Gly Gln Leu Glu Gln Arg

ACC AAA TAT AAA AGA GAG CCA GTA TCA TTG ACC CTG GCC CTT CTA CTA GGA GGA TTA ACC   1380
Thr Lys Tyr Lys Arg Glu Pro Val Ser Leu Thr Leu Ala Leu Leu Leu Gly Gly Leu Thr

ATG GGA GGG ATT GCA GCT GGA ATA GGG ACG GGG ACC ACT GCC TTA ATT AAA ACC CAG CAG   1440
Met Gly Gly Ile Ala Ala Gly Ile Gly Thr Gly Thr Thr Ala Leu Ile Lys Thr Gln Gln

TTT GAG CAG CTT CAT GCC GCT ATC CAG ACA GAC CTC AAC GAA GTC GAA AAG TCA ATT ACC   1500
Phe Glu Gln Leu His Ala Ala Ile Gln Thr Asp Leu Asn Glu Val Glu Lys Ser Ile Thr

AAC CTA GAA AAG TCA CTG ACC TCG TTG TCT GAA GTA GTC CTA CAG AAC CGC AGA GGC CTA   1560
Asn Leu Glu Lys Ser Leu Thr Ser Leu Ser Glu Val Val Leu Gln Asn Arg Arg Gly Leu

GAT TTG CTA TTC CTA AAG GAG GGA GGT CTC TGC GCA GCC CTA AAA GAA GAA TGT TGT TTT   1620
Asp Leu Leu Phe Leu Lys Glu Gly Gly Leu Cys Ala Ala Leu Lys Glu Glu Cys Cys Phe

TAT GCA GAC CAC ACG GGG CTA GTG AGA GAC AGC ATG GCC AAA TTA AGA GAA AGG CTT AAT   1680
Tyr Ala Asp His Thr Gly Leu Val Arg Asp Ser Met Ala Lys Leu Arg Glu Arg Leu Asn

CAG AGA CAA AAA CTA TTT GAG ACA GGC CAA GGA TGG TTC GAA GGG CTG TTT AAT AGA TCC   1740
Gln Arg Gln Lys Leu Phe Glu Thr Gly Gln Gly Trp Phe Glu Gly Leu Phe Asn Arg Ser

CCC TGG TTT ACC ACC TTA ATC TCC ACC ATC ATG GGA CCT CTA ATA GTA CTC TTA CTG ATC   1800
Pro Trp Phe Thr Thr Leu Ile Ser Thr Ile Met Gly Pro Leu Ile Val Leu Leu Leu Ile

TTA CTC TTT GGA CCT TGC ATT CTC AAT CGA TTG GTC CAA TTT GTT AAA GAC AGG ATC TCA   1860
Leu Leu Phe Gly Pro Cys Ile Leu Asn Arg Leu Val Gln Phe Val Lys Asp Arg Ile Ser

GTG GTC CAG GCT CTG GTT TTG ACT CAG CAA TAT CAC CAG CTA AAA CCC ATA GAG TAC GAG   1920
Val Val Gln Ala Leu Val Leu Thr Gln Gln Tyr His Gln Leu Lys Pro Ile Glu Tyr Glu

CCA TGA                                                                           1926
Pro ...
```

FIG. 18 (continued)

AdelHRless Sequence

```
          10          20          30          40          50          60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
ATG GCG CGT TCA ACG CTC TCA AAA CCC CCT CAA GAT AAG ATT AAC CCG TGG AAG CCC TTA     60
Met Ala Arg Ser Thr Leu Ser Lys Pro Pro Gln Asp Lys Ile Asn Pro Trp Lys Pro Leu

ATA GTC ATG GGA GTC CTG TTA GGA GTA GGG ATG GCA GAG AGC CCC CAG GTC TTT AAT GTA    120
Ile Val Met Gly Val Leu Leu Gly Val Gly Met Ala Glu Ser Pro Gln Val Phe Asn Val

ACC TGG AGA GTC ACC AAC CTG ATG ACT GGG CGT ACC GCC AAT GCC ACC TCC CTC CTG GGA    180
Thr Trp Arg Val Thr Asn Leu Met Thr Gly Arg Thr Ala Asn Ala Thr Ser Leu Leu Gly

ACT GTA CAA GAT GCC TTC CCA AAA TTA TAT TTT GAT CTA TGT GAT CTG GTC GGA GAG GAG    240
Thr Val Gln Asp Ala Phe Pro Lys Leu Tyr Phe Asp Leu Cys Asp Leu Val Gly Glu Glu

TGG GAC CCT TCA GAC CAG GAA CCG TAT GTC GGG TAT GGC TGC AAG TAC CCC GCA GGG AGA    300
Trp Asp Pro Ser Asp Gln Glu Pro Tyr Val Gly Tyr Gly Cys Lys Tyr Pro Ala Gly Arg

CAG CGG ACC CGG ACT TTT GAC TTT TAC GTG TGC CCT GGG CAT ACC GTA AAG TCG GGG TGT    360
Gln Arg Thr Arg Thr Phe Asp Phe Tyr Val Cys Pro Gly His Thr Val Lys Ser Gly Cys

GGG GGA CCA GGA GAG GGC TAC TGT GGT AAA TGG GGG TGT GAA ACC ACC GGA CAG GCT TAC    420
Gly Gly Pro Gly Glu Gly Tyr Cys Gly Lys Trp Gly Cys Glu Thr Thr Gly Gln Ala Tyr

TGG AAG CCC ACA TCA TCG TGG GAC CTA ATC TCC CTT AAG CGC GGT AAC ACC CCC TGG GAC    480
Trp Lys Pro Thr Ser Ser Trp Asp Leu Ile Ser Leu Lys Arg Gly Asn Thr Pro Trp Asp

ACG GGA TGC TCT AAA GTT GCC TGT GGC CCC TGC TAC GAC CTC TCC AAA GTA TCC AAT TCC    540
Thr Gly Cys Ser Lys Val Ala Cys Gly Pro Cys Tyr Asp Leu Ser Lys Val Ser Asn Ser

TTC CAA GGG GCT ACT CGA GGG GGC AGA TGC AAC CCT CTA GTC CTA GAA TTC ACT GAT GCA    600
Phe Gln Gly Ala Thr Arg Gly Gly Arg Cys Asn Pro Leu Val Leu Glu Phe Thr Asp Ala

GGA AAA AAG GCT AAC TGG GAC GGG CCC AAA TCG TGG GGA CTG AGA CTG TAC CGG ACA GGA    660
Gly Lys Lys Ala Asn Trp Asp Gly Pro Lys Ser Trp Gly Leu Arg Leu Tyr Arg Thr Gly

ACA GAT CCT ATT ACC ATG TTC TCC CTG ACC CGG CAG GTC CTT AAT GTG GGA CCC CGA GTC    720
Thr Asp Pro Ile Thr Met Phe Ser Leu Thr Arg Gln Val Leu Asn Val Gly Pro Arg Val

CCC ATA GGG CCC AAC CCA GTA TTA CCC GAC CAA AGA CTC CCT TCC TCA CCA ATA GAG ATT    780
Pro Ile Gly Pro Asn Pro Val Leu Pro Asp Gln Arg Leu Pro Ser Ser Pro Ile Glu Ile

GTA CCG GCT CCA CAG CCA CCT AGC CCC CTC AAT ACC AGT TAC CCC CCT TCC ACT ACC AGT    840
Val Pro Ala Pro Gln Pro Pro Ser Pro Leu Asn Thr Ser Tyr Pro Pro Ser Thr Thr Ser

ACA CCC TCA ACC TCC CCT ACA AGT CCA AGT GTC CCA CAG CCA CCC CCA GGA ACT GGA GAT    900
Thr Pro Ser Thr Ser Pro Thr Ser Pro Ser Val Pro Gln Pro Pro Pro Gly Thr Gly Asp

AGA CTA CTA GCT CTA GTC AAA GGA GCC TAT CAG GCG CTT AAC CTC ACC AAT CCC GAC AAG    960
Arg Leu Leu Ala Leu Val Lys Gly Ala Tyr Gln Ala Leu Asn Leu Thr Asn Pro Asp Lys

ACC CAA GAA TGT TGG CTG TGC TTA GTG TCG GGA CCT CCT TAT TAC GAA GGA GTA GCG GTC   1020
Thr Gln Glu Cys Trp Leu Cys Leu Val Ser Gly Pro Pro Tyr Tyr Glu Gly Val Ala Val

GTG GGC ACT TAT ACC AAT CAT TCC ACC GCT CCG GCC AAC TGT ACG GCC ACT TCC CAA CAT   1080
Val Gly Thr Tyr Thr Asn His Ser Thr Ala Pro Ala Asn Cys Thr Ala Thr Ser Gln His

AAG CTT ACC CTA TCT GAA GTG ACA GGA CAG GGC CTA TGC ATG GGG GCA GTA CCT AAA ACT   1140
Lys Leu Thr Leu Ser Glu Val Thr Gly Gln Gly Leu Cys Met Gly Ala Val Pro Lys Thr

CAC CAG GCC TTA TGT AAC ACC ACC CAA AGC GCC GGC TCA GGA TCC TAC TAC CTT GCA GCA   1200
His Gln Ala Leu Cys Asn Thr Thr Gln Ser Ala Gly Ser Gly Ser Tyr Tyr Leu Ala Ala
```

FIG. 19

AdelHRless Sequence

```
           10           20           30            40           50           60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
CCC GCC GGA ACA ATG TGG GCT TGC AGC ACT GGA TTG ACT CCC TGC TTG TCC ACC ACG GTG   1260
Pro Ala Gly Thr Met Trp Ala Cys Ser Thr Gly Leu Thr Pro Cys Leu Ser Thr Thr Val

CTC AAT CTA ACC ACA GAT TAT TGT GTA TTA GTT GAA CTC TGG CCC AGA GTA ATT TAC CAC   1320
Leu Asn Leu Thr Thr Asp Tyr Cys Val Leu Val Glu Leu Trp Pro Arg Val Ile Tyr His

TCC CCC GAT TAT ATG TAT GGT CAG CTT GAA CAG CGT ACC AAA TAT AAA AGA GAG CCA GTA   1380
Ser Pro Asp Tyr Met Tyr Gly Gln Leu Glu Gln Arg Thr Lys Tyr Lys Arg Glu Pro Val

TCA TTG ACC CTG GCC CTT CTA CTA GGA GGA TTA ACC ATG GGA GGG ATT GCA GCT GGA ATA   1440
Ser Leu Thr Leu Ala Leu Leu Leu Gly Gly Leu Thr Met Gly Gly Ile Ala Ala Gly Ile

GGG ACG GGG ACC ACT GCC TTA ATT AAA ACC CAG CAG TTT GAG CAG CTT CAT GCC GCT ATC   1500
Gly Thr Gly Thr Thr Ala Leu Ile Lys Thr Gln Gln Phe Glu Gln Leu His Ala Ala Ile

CAG ACA GAC CTC AAC GAA GTC GAA AAG TCA ATT ACC AAC CTA GAA AAG TCA CTG ACC TCG   1560
Gln Thr Asp Leu Asn Glu Val Glu Lys Ser Ile Thr Asn Leu Glu Lys Ser Leu Thr Ser

TTG TCT GAA GTA GTC CTA CAG AAC CGC AGA GGC CTA GAT TTG CTA TTC CTA AAG GAG GGA   1620
Leu Ser Glu Val Val Leu Gln Asn Arg Arg Gly Leu Asp Leu Leu Phe Leu Lys Glu Gly

GGT CTC TGC GCA GCC CTA AAA GAA GAA TGT TGT TTT TAT GCA GAC CAC ACG GGG CTA GTG   1680
Gly Leu Cys Ala Ala Leu Lys Glu Glu Cys Cys Phe Tyr Ala Asp His Thr Gly Leu Val

AGA GAC AGC ATG GCC AAA TTA AGA GAA AGG CTT AAT CAG AGA CAA AAA CTA TTT GAG ACA   1740
Arg Asp Ser Met Ala Lys Leu Arg Glu Arg Leu Asn Gln Arg Gln Lys Leu Phe Glu Thr

GGC CAA GGA TGG TTC GAA GGG CTG TTT AAT AGA TCC CCC TGG TTT ACC ACC TTA ATC TCC   1800
Gly Gln Gly Trp Phe Glu Gly Leu Phe Asn Arg Ser Pro Trp Phe Thr Thr Leu Ile Ser

ACC ATC ATG GGA CCT CTA ATA GTA CTC TTA CTG ATC TTA CTC TTT GGA CCT TGC ATT CTC   1860
Thr Ile Met Gly Pro Leu Ile Val Leu Leu Leu Ile Leu Leu Phe Gly Pro Cys Ile Leu

AAT CGA TTA GTT CAA TTT GTT AAA GAC AGG ATC TCA GTA GTC CAG GCT TTA               1911
Asn Arg Leu Val Gln Phe Val Lys Asp Arg Ile Ser Val Val Gln Ala Leu
```

FIG. 19 (continued)

PROMOdelHRless Sequence

```
                  10                  20                  30                  40                  50                  60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
ATG GCG CGT TCA ACG CTC TCA AAA CCC CCT CAA GAT AAG ATT AAC CCG TGG AAG CCC TTA     60
Met Ala Arg Ser Thr Leu Ser Lys Pro Pro Gln Asp Lys Ile Asn Pro Trp Lys Pro Leu

ATA GTC ATG GGA GTC CTG TTA GGA GTA GGG ATG GCA GAG AGC CCC CAG GTC TTT AAT GTA    120
Ile Val Met Gly Val Leu Leu Gly Val Gly Met Ala Glu Ser Pro Gln Val Phe Asn Val

ACC TGG AGA GTC ACC AAC CTG ATG ACT GGG CGT ACC GCC AAT GCC ACC TCC CTC CTG GGA    180
Thr Trp Arg Val Thr Asn Leu Met Thr Gly Arg Thr Ala Asn Ala Thr Ser Leu Leu Gly

ACT GTA CAA GAT GCC TTC CCA AAA TTA TAT TTT GAT CTA TGT GAT CTG GTC GGA GAG GAG    240
Thr Val Gln Asp Ala Phe Pro Lys Leu Tyr Phe Asp Leu Cys Asp Leu Val Gly Glu Glu

TGG GAC CCT TCA GAC CAG GAA CCG TAT GTC GGG TAT GGC TGC AAG TAC CCC GCA GGG AGA    300
Trp Asp Pro Ser Asp Gln Glu Pro Tyr Val Gly Tyr Gly Cys Lys Tyr Pro Ala Gly Arg

CAG CGG ACC CGG ACT TTT GAC TTT TAC GTG TGC CCT GGG CAT ACC GTA AAG TCG GGG TGT    360
Gln Arg Thr Arg Thr Phe Asp Phe Tyr Val Cys Pro Gly His Thr Val Lys Ser Gly Cys

GGG GGA CCA GGA GAG GGC TAC TGT GGT AAA TGG GGG TGT GAA ACC ACC GGA CAG GCT TAC    420
Gly Gly Pro Gly Glu Gly Tyr Cys Gly Lys Trp Gly Cys Glu Thr Thr Gly Gln Ala Tyr

TGG AAG CCC ACA TCA TCG TGG GAC CTA ATC TCC CTT AAG CGC GGT AAC ACC CCC TGG GAC    480
Trp Lys Pro Thr Ser Ser Trp Asp Leu Ile Ser Leu Lys Arg Gly Asn Thr Pro Trp Asp

ACG GGA TGC TCT AAA GTT GCC TGT GGC CCC TGC TAC GAC CTC TCC AAA GTA TCC AAT TCC    540
Thr Gly Cys Ser Lys Val Ala Cys Gly Pro Cys Tyr Asp Leu Ser Lys Val Ser Asn Ser

TTC CAA GGG GCT ACT CGA GGG GGC AGA TGC AAC CCT CTA GTC CTA GAA TTC ACT GAT GCA    600
Phe Gln Gly Ala Thr Arg Gly Gly Arg Cys Asn Pro Leu Val Leu Glu Phe Thr Asp Ala

GGA AAA AAG GCT AAC TGG GAC GGG CCC AAA TCG TGG GGA CTG AGA CTG TAC CGG ACA GGA    660
Gly Lys Lys Ala Asn Trp Asp Gly Pro Lys Ser Trp Gly Leu Arg Leu Tyr Arg Thr Gly

ACA GAT CCT ATT ACC ATG TTC TCC CTG ACC CGG CAG GTC CTT AAT GTG GGA CCC CGA GTC    720
Thr Asp Pro Ile Thr Met Phe Ser Leu Thr Arg Gln Val Leu Asn Val Gly Pro Arg Val

CCC ATA GGG CCC AAC CCC GTT CTG GCA GAC CAA CAG CCA CTC TCC AAG CCC AAA CCT GTT    780
Pro Ile Gly Pro Asn Pro Val Leu Ala Asp Gln Gln Pro Leu Ser Lys Pro Lys Pro Val

AAG TCG CCT TCA GTC ACC AAA CCA CCC AGT GGG ACT CCT CTC TCC CCT ACC CAA CTT CCA    840
Lys Ser Pro Ser Val Thr Lys Pro Pro Ser Gly Thr Pro Leu Ser Pro Thr Gln Leu Pro

CCG GCG GGT ACC GGA GAT AGA CTA CTA GCT CTA GTC AAA GGA GCC TAT CAG GCG CTT AAC    900
Pro Ala Gly Thr Gly Asp Arg Leu Leu Ala Leu Val Lys Gly Ala Tyr Gln Ala Leu Asn

CTC ACC AAT CCC GAC AAG ACC CAA GAA TGT TGG CTG TGC TTA GTG TCG GGA CCT CCT TAT    960
Leu Thr Asn Pro Asp Lys Thr Gln Glu Cys Trp Leu Cys Leu Val Ser Gly Pro Pro Tyr

TAC GAA GGA GTA GCG GTC GTG GGC ACT TAT ACC AAT CAT TCC ACC GCT CCG GCC AAC TGT   1020
Tyr Glu Gly Val Ala Val Val Gly Thr Tyr Thr Asn His Ser Thr Ala Pro Ala Asn Cys

ACG GCC ACT TCC CAA CAT AAG CTT ACC CTA TCT GAA GTG ACA GGA CAG GGC CTA TGC ATG   1080
Thr Ala Thr Ser Gln His Lys Leu Thr Leu Ser Glu Val Thr Gly Gln Gly Leu Cys Met

GGG GCA GTA CCT AAA ACT CAC CAG GCC TTA TGT AAC ACC ACC CAA AGC GCC GGC TCA GGA   1140
Gly Ala Val Pro Lys Thr His Gln Ala Leu Cys Asn Thr Thr Gln Ser Ala Gly Ser Gly

TCC TAC TAC CTT GCA GCA CCC GCC GGA ACA ATG TGG GCT TGC AGC ACT GGA TTG ACT CCC   1200
Ser Tyr Tyr Leu Ala Ala Pro Ala Gly Thr Met Trp Ala Cys Ser Thr Gly Leu Thr Pro
```

FIG. 20

PROMOdelHRless Sequence

```
                10                  20                  30                  40                  50                  60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
TGC TTG TCC ACC ACG GTG CTC AAT CTA ACC ACA GAT TAT TGT GTA TTA GTT GAA CTC TGG    1260
Cys Leu Ser Thr Thr Val Leu Asn Leu Thr Thr Asp Tyr Cys Val Leu Val Glu Leu Trp

CCC AGA GTA ATT TAC CAC TCC CCC GAT TAT ATG TAT GGT CAG CTT GAA CAG CGT ACC AAA    1320
Pro Arg Val Ile Tyr His Ser Pro Asp Tyr Met Tyr Gly Gln Leu Glu Gln Arg Thr Lys

TAT AAA AGA GAG CCA GTA TCA TTG ACC CTG GCC CTT CTA CTA GGA GGA TTA ACC ATG GGA    1380
Tyr Lys Arg Glu Pro Val Ser Leu Thr Leu Ala Leu Leu Leu Gly Gly Leu Thr Met Gly

GGG ATT GCA GCT GGA ATA GGG ACG GGG ACC ACT GCC TTA ATT AAA ACC CAG CAG TTT GAG    1440
Gly Ile Ala Ala Gly Ile Gly Thr Gly Thr Thr Ala Leu Ile Lys Thr Gln Gln Phe Glu

CAG CTT CAT GCC GCT ATC CAG ACA GAC CTC AAC GAA GTC GAA AAG TCA ATT ACC AAC CTA    1500
Gln Leu His Ala Ala Ile Gln Thr Asp Leu Asn Glu Val Glu Lys Ser Ile Thr Asn Leu

GAA AAG TCA CTG ACC TCG TTG TCT GAA GTA GTC CTA CAG AAC CGC AGA GGC CTA GAT TTG    1560
Glu Lys Ser Leu Thr Ser Leu Ser Glu Val Val Leu Gln Asn Arg Arg Gly Leu Asp Leu

CTA TTC CTA AAG GAG GGA GGT CTC TGC GCA GCC CTA AAA GAA GAA TGT TGT TTT TAT GCA    1620
Leu Phe Leu Lys Glu Gly Gly Leu Cys Ala Ala Leu Lys Glu Glu Cys Cys Phe Tyr Ala

GAC CAC ACG GGG CTA GTG AGA GAC AGC ATG GCC AAA TTA AGA GAA AGG CTT AAT CAG AGA    1680
Asp His Thr Gly Leu Val Arg Asp Ser Met Ala Lys Leu Arg Glu Arg Leu Asn Gln Arg

CAA AAA CTA TTT GAG ACA GGC CAA GGA TGG TTC GAA GGG CTG TTT AAT AGA TCC CCC TGG    1740
Gln Lys Leu Phe Glu Thr Gly Gln Gly Trp Phe Glu Gly Leu Phe Asn Arg Ser Pro Trp

TTT ACC ACC TTA ATC TCC ACC ATC ATG GGA CCT CTA ATA GTA CTC TTA CTG ATC TTA CTC    1800
Phe Thr Thr Leu Ile Ser Thr Ile Met Gly Pro Leu Ile Val Leu Leu Leu Ile Leu Leu

TTT GGA CCT TGC ATT CTC AAT CGA TTA GTT CAA TTT GTT AAA GAC AGG ATC TCA GTA GTC    1860
Phe Gly Pro Cys Ile Leu Asn Arg Leu Val Gln Phe Val Lys Asp Arg Ile Ser Val Val

CAG GCT TTA                                                                        1869
Gln Ala Leu
```

FIG. 20 (continued)

GALVdelHRless Sequence

```
            10          20          30          40          50          60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
ATG GTA TTG CTG CCT GGG TCC ATG CTT CTC ACC TCA AAC CTG CAC CAC CTT CGG CAC CAG      60
Met Val Leu Leu Pro Gly Ser Met Leu Leu Thr Ser Asn Leu His His Leu Arg His Gln

ATG AGT CCT GGG AGC TGG AAA AGA CTG ATC ATC CTC TTA AGC TGC GTA TTC GGC GGC GGC     120
Met Ser Pro Gly Ser Trp Lys Arg Leu Ile Ile Leu Leu Ser Cys Val Phe Gly Gly Gly

GGG ACG AGT CTG CAG AAT AAG AAC CCC CAG CCC ATG ACC CTC ACT TGG CAG GTA CTG TCC     180
Gly Thr Ser Leu Gln Asn Lys Asn Pro Gln Pro Met Thr Leu Thr Trp Gln Val Leu Ser

CAA ACT GGA GAC GTT GTC TGG GAT ACA AAG GCA GTC CAG CCC CCT TGG ACT TGG TGG CCC     240
Gln Thr Gly Asp Val Val Trp Asp Thr Lys Ala Val Gln Pro Pro Trp Thr Trp Trp Pro

ACA CTT AAA CCT GAT GTA TGT GCC TTG GCG GCT AGT CTT GAG TCC TGG GAT ATC CCG GGA     300
Thr Leu Lys Pro Asp Val Cys Ala Leu Ala Ala Ser Leu Glu Ser Trp Asp Ile Pro Gly

ACC GAT GTC TCG TCC TCT AAA CGA GTC AGA CCT CCG GAC TCA GAC TAT ACT GCC GCT TAT     360
Thr Asp Val Ser Ser Ser Lys Arg Val Arg Pro Pro Asp Ser Asp Tyr Thr Ala Ala Tyr

AAG CAA ATC ACC TGG GGA GCC ATA GGG TGC AGC TAC CCT CGG GCT AGG ACT AGA ATG GCA     420
Lys Gln Ile Thr Trp Gly Ala Ile Gly Cys Ser Tyr Pro Arg Ala Arg Thr Arg Met Ala

AGC TCT ACC TTC TAC GTA TGT CCC CGG GAT GGC CGG ACC CTT TCA GAA GCT AGA AGG TGC     480
Ser Ser Thr Phe Tyr Val Cys Pro Arg Asp Gly Arg Thr Leu Ser Glu Ala Arg Arg Cys

GGG GGG CTA GAA TCC CTA TAC TGT AAA GAA TGG GAT TGT GAG ACC ACG GGG ACC GGT TAT     540
Gly Gly Leu Glu Ser Leu Tyr Cys Lys Glu Trp Asp Cys Glu Thr Thr Gly Thr Gly Tyr

TGG CTA TCT AAA TCC TCA AAA GAC CTC ATA ACT GTA AAA TGG GAC CAA AAT AGC GAA TGG     600
Trp Leu Ser Lys Ser Ser Lys Asp Leu Ile Thr Val Lys Trp Asp Gln Asn Ser Glu Trp

ACT CAA AAA TTT CAA CAG TGT CAC CAG ACC GGC TGG TGT AAC CCC CTT AAA ATA GAT TTC     660
Thr Gln Lys Phe Gln Gln Cys His Gln Thr Gly Trp Cys Asn Pro Leu Lys Ile Asp Phe

ACA GAC AAA GGA AAA TTA TCC AAG GAC TGG ATA ACG GGA AAA ACC TGG GGA TTA AGA TTC     720
Thr Asp Lys Gly Lys Leu Ser Lys Asp Trp Ile Thr Gly Lys Thr Trp Gly Leu Arg Phe

TAT GTG TCT GGA CAT CCA GGC GTA CAG TTC ACC ATT CGC TTA AAA ATC ACC AAC ATG CCA     780
Tyr Val Ser Gly His Pro Gly Val Gln Phe Thr Ile Arg Leu Lys Ile Thr Asn Met Pro

GCT GTG GCA GTA GGT CCT GAC CTC GTC CTT GTG GAA CAA GGA CCT CCT AGA ACG TCC CTC     840
Ala Val Ala Val Gly Pro Asp Leu Val Leu Val Glu Gln Gly Pro Pro Arg Thr Ser Leu

GCT CTC CCA CCT CCT CTT CCC CCA AGG GAA GCG CCA CCG CCA TCT CTC CCC GAC TCT AAC     900
Ala Leu Pro Pro Pro Leu Pro Pro Arg Glu Ala Pro Pro Pro Ser Leu Pro Asp Ser Asn

TCC ACA GCC CTG GCG ACT AGT GCA CAA ACT CCC ACG GTG AGA AAA ACA ATT GTT ACC CTA     960
Ser Thr Ala Leu Ala Thr Ser Ala Gln Thr Pro Thr Val Arg Lys Thr Ile Val Thr Leu

AAC ACT CCG CCT CCC ACC ACA GGC GAC AGA CTT TTT GAT CTT GTG CAG GGG GCC TTC CTA    1020
Asn Thr Pro Pro Pro Thr Thr Gly Asp Arg Leu Phe Asp Leu Val Gln Gly Ala Phe Leu

ACC TTA AAT GCT ACC AAC CCA GGG GCC ACT GAG TCT TGC TGG CTT TGT TTG GCC ATG GGC    1080
Thr Leu Asn Ala Thr Asn Pro Gly Ala Thr Glu Ser Cys Trp Leu Cys Leu Ala Met Gly

CCC CCT TAT TAT GAA GCA ATA GCC TCA TCA GGA GAG GTC GCC TAC TCC ACC GAC CTT GAC    1140
Pro Pro Tyr Tyr Glu Ala Ile Ala Ser Ser Gly Glu Val Ala Tyr Ser Thr Asp Leu Asp

CGG TGC CGC TGG GGG ACC CAA GGA AAG CTC ACC CTC ACT GAG GTC TCA GGA CAC GGG TTG    1200
Arg Cys Arg Trp Gly Thr Gln Gly Lys Leu Thr Leu Thr Glu Val Ser Gly His Gly Leu
```

FIG. 21

GALVdelHRless Sequence

```
                10              20              30              40              50              60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
TGC ATA GGA AAG GTG CCC TTT ACC CAT CAG CAT CTC TGC AAT CAG ACC CTA TCC ATC AAT   1260
Cys Ile Gly Lys Val Pro Phe Thr His Gln His Leu Cys Asn Gln Thr Leu Ser Ile Asn

TCC TCC GGA GAC CAT CAG TAT CTG CTC CCC TCC AAC CAT AGC TGG TGG GCT TGC AGC ACT   1320
Ser Ser Gly Asp His Gln Tyr Leu Leu Pro Ser Asn His Ser Trp Trp Ala Cys Ser Thr

GGC CTC ACC CCT TGC CTC TCC ACC TCA GTT TTT AAT CAG ACT AGA GAT TTC TGT ATC CAG   1380
Gly Leu Thr Pro Cys Leu Ser Thr Ser Val Phe Asn Gln Thr Arg Asp Phe Cys Ile Gln

GTC CAG CTG ATT CCT CGC ATC TAT TAC TAT CCT GAA GAA GTT TTG TTA CAG GCC TAT GAC   1440
Val Gln Leu Ile Pro Arg Ile Tyr Tyr Tyr Pro Glu Glu Val Leu Leu Gln Ala Tyr Asp

AAT TCT CAC CCC AGG ACT AAA AGA GAG GCT GTC TCA CTT ACC CTA GCT GTT TTA CTG GGG   1500
Asn Ser His Pro Arg Thr Lys Arg Glu Ala Val Ser Leu Thr Leu Ala Val Leu Leu Gly

TTG GGA ATC ACG GCG GGA ATA GGT ACT GGT TCA ACT GCC TTA ATT AAA GGA CCT ATA GAC   1560
Leu Gly Ile Thr Ala Gly Ile Gly Thr Gly Ser Thr Ala Leu Ile Lys Gly Pro Ile Asp

CTC CAG CAA GGC CTG ACA AGC CTC CAG ATC GCC ATA GAT GCT GAC CTC CGG GCC CTC CAA   1620
Leu Gln Gln Gly Leu Thr Ser Leu Gln Ile Ala Ile Asp Ala Asp Leu Arg Ala Leu Gln

GAC TCA GTC AGC AAG TTA GAG GAC TCA CTG ACT TCC CTG TCC GAG GTA GTG CTC CAA AAT   1680
Asp Ser Val Ser Lys Leu Glu Asp Ser Leu Thr Ser Leu Ser Glu Val Val Leu Gln Asn

AGG AGA GGC CTT GAC TTG CTG TTT CTA AAA GAA GGT GGC CTC TGT GCG GCC CTA AAG GAA   1740
Arg Arg Gly Leu Asp Leu Leu Phe Leu Lys Glu Gly Gly Leu Cys Ala Ala Leu Lys Glu

GAG TGC TGT TTT TAC ATA GAC CAC TCA GGT GCA GTA CGG GAC TCC ATG AAA AAA CTC AAA   1800
Glu Cys Cys Phe Tyr Ile Asp His Ser Gly Ala Val Arg Asp Ser Met Lys Lys Leu Lys

GAA AAA CTG GAT AAA AGA CAG TTA GAG CGC CAG AAA AGC CAA AAC TGG TAT GAA GGA TGG   1860
Glu Lys Leu Asp Lys Arg Gln Leu Glu Arg Gln Lys Ser Gln Asn Trp Tyr Glu Gly Trp

TTC AAT AAC TCC CCT TGG TTC ACT ACC CTG CTA TCA ACC ATC GCT GGG CCC CTA TTA CTC   1920
Phe Asn Asn Ser Pro Trp Phe Thr Thr Leu Leu Ser Thr Ile Ala Gly Pro Leu Leu Leu

CTC CTT CTG TTG CTC ATC CTC GGG CCA TGC ATC ATC AAT AAG TTA GTT CAA TTC ATC AAT   1980
Leu Leu Leu Leu Leu Ile Leu Gly Pro Cys Ile Ile Asn Lys Leu Val Gln Phe Ile Asn

GAT AGG ATA AGT GCA TGT                                                           1998
Asp Arg Ile Ser Ala Cys
```

FIG. 21 (continued)

C2delHRless Sequence

```
                 10                  20                  30                  40                  50                  60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890

ATG GCG CGT TCA ACG CTC TCA AAA CCC CCT CAA GAT AAG ATT AAC CCG TGG AAG CCC TTA     60
Met Ala Arg Ser Thr Leu Ser Lys Pro Pro Gln Asp Lys Ile Asn Pro Trp Lys Pro Leu

ATA GTC ATG GGA GTC CTG TTA GGA GTA GGG ATG GCA GAG AGC CCC CAG GTC TTT AAT GTA    120
Ile Val Met Gly Val Leu Leu Gly Val Gly Met Ala Glu Ser Pro Gln Val Phe Asn Val

ACC TGG AGA GTC ACC AAC CTG ATG ACT GGG CGT ACC GCC AAT GCC ACC TCC CTC CTG GGA    180
Thr Trp Arg Val Thr Asn Leu Met Thr Gly Arg Thr Ala Asn Ala Thr Ser Leu Leu Gly

ACT GTA CAA GAT GCC TTC CCA AAA TTA TAT TTT GAT CTA TGT GAT CTG GTC GGA GAG GAG    240
Thr Val Gln Asp Ala Phe Pro Lys Leu Tyr Phe Asp Leu Cys Asp Leu Val Gly Glu Glu

TGG GAC CCT TCA GAC CAG GAA CCG TAT GTC GGG TAT GGC TGC AAG TAC CCC GCA GGG AGA    300
Trp Asp Pro Ser Asp Gln Glu Pro Tyr Val Gly Tyr Gly Cys Lys Tyr Pro Ala Gly Arg

CAG CGG ACC CGG ACT TTT GAC TTT TAC GTG TGC CCT GGG CAT ACC GTA AAG TCG GGG TGT    360
Gln Arg Thr Arg Thr Phe Asp Phe Tyr Val Cys Pro Gly His Thr Val Lys Ser Gly Cys

GGG GGA CCA GGA GAG GGC TAC TGT GGT AAA TGG GGG TGT GAA ACC ACC GGA CAG GCT TAC    420
Gly Gly Pro Gly Glu Gly Tyr Cys Gly Lys Trp Gly Cys Glu Thr Thr Gly Gln Ala Tyr

TGG AAG CCC ACA TCA TCG TGG GAC CTA ATC TCC CTT AAG CGC GGT AAC ACC CCC TGG GAC    480
Trp Lys Pro Thr Ser Ser Trp Asp Leu Ile Ser Leu Lys Arg Gly Asn Thr Pro Trp Asp

ACG GGA TGC TCT AAA GTT GCC TGT GGC CCC TGC TAC GAC CTC TCC AAA GTA TCC AAT TCC    540
Thr Gly Cys Ser Lys Val Ala Cys Gly Pro Cys Tyr Asp Leu Ser Lys Val Ser Asn Ser

TTC CAA GGG GCT ACT CGA GGG GGC AGA TGC AAC CCT CTA GTC CTA GAA TTC ACT GAT GCA    600
Phe Gln Gly Ala Thr Arg Gly Gly Arg Cys Asn Pro Leu Val Leu Glu Phe Thr Asp Ala

GGA AAA AAG GCT AAC TGG GAC GGG CCC AAA TCG TGG GGA CTG AGA CTG TAC CGG ACA GGA    660
Gly Lys Lys Ala Asn Trp Asp Gly Pro Lys Ser Trp Gly Leu Arg Leu Tyr Arg Thr Gly

ACA GAT CCT ATT ACC ATG TTC TCC CTG ACC CGG CAG GTC CTT AAT GTG GGA CCC CGA GTC    720
Thr Asp Pro Ile Thr Met Phe Ser Leu Thr Arg Gln Val Leu Asn Val Gly Pro Arg Val

CCC ATA GGG CCC AAC CCA GTA TTA CCC GAC CAA AGA CTC CCT TCC TCA CCA ATA GAG ATT    780
Pro Ile Gly Pro Asn Pro Val Leu Pro Asp Gln Arg Leu Pro Ser Ser Pro Ile Glu Ile

GTA CCG GCT CCA CAG CCA CCT AGC CCC CTC AAT ACC AGT TAC CCC CCT TCC ACT ACC AGT    840
Val Pro Ala Pro Gln Pro Pro Ser Pro Leu Asn Thr Ser Tyr Pro Pro Ser Thr Thr Ser

ACA CCC CCC CCA GGT ACC GGA GAT AGA CTA CTA GCT CTA GTC AAA GGA GCC TAT CAG GCG    900
Thr Pro Pro Pro Gly Thr Gly Asp Arg Leu Leu Ala Leu Val Lys Gly Ala Tyr Gln Ala

CTT AAC CTC ACC AAT CCC GAC AAG ACC CAA GAA TGT TGG CTG TGC TTA GTG TCG GGA CCT    960
Leu Asn Leu Thr Asn Pro Asp Lys Thr Gln Glu Cys Trp Leu Cys Leu Val Ser Gly Pro

CCT TAT TAC GAA GGA GTA GCG GTC GTG GGC ACT TAT ACC AAT CAT TCC ACC GCT CCG GCC   1020
Pro Tyr Tyr Glu Gly Val Ala Val Val Gly Thr Tyr Thr Asn His Ser Thr Ala Pro Ala

AAC TGT ACG GCC ACT TCC CAA CAT AAG CTT ACC CTA TCT GAA GTG ACA GGA CAG GGC CTA   1080
Asn Cys Thr Ala Thr Ser Gln His Lys Leu Thr Leu Ser Glu Val Thr Gly Gln Gly Leu

TGC ATG GGG GCA GTA CCT AAA ACT CAC CAG GCC TTA TGT AAC ACC ACC CAA AGC GCC GGC   1140
Cys Met Gly Ala Val Pro Lys Thr His Gln Ala Leu Cys Asn Thr Thr Gln Ser Ala Gly

TCA GGA TCC TAC TAC CTT GCA GCA CCC GCC GGA ACA ATG TGG GCT TGC AGC ACT GGA TTG   1200
Ser Gly Ser Tyr Tyr Leu Ala Ala Pro Ala Gly Thr Met Trp Ala Cys Ser Thr Gly Leu
```

FIG. 22

C2delHRless Sequence

```
             10          20          30          40          50          60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
ACT CCC TGC TTG TCC ACC ACG GTG CTC AAT CTA ACC ACA GAT TAT TGT GTA TTA GTT GAA   1260
Thr Pro Cys Leu Ser Thr Thr Val Leu Asn Leu Thr Thr Asp Tyr Cys Val Leu Val Glu

CTC TGG CCC AGA GTA ATT TAC CAC TCC CCC GAT TAT ATG TAT GGT CAG CTT GAA CAG CGT   1320
Leu Trp Pro Arg Val Ile Tyr His Ser Pro Asp Tyr Met Tyr Gly Gln Leu Glu Gln Arg

ACC AAA TAT AAA AGA GAG CCA GTA TCA TTG ACC CTG GCC CTT CTA CTA GGA GGA TTA ACC   1380
Thr Lys Tyr Lys Arg Glu Pro Val Ser Leu Thr Leu Ala Leu Leu Leu Gly Gly Leu Thr

ATG GGA GGG ATT GCA GCT GGA ATA GGG ACG GGG ACC ACT GCC TTA ATT AAA ACC CAG CAG   1440
Met Gly Gly Ile Ala Ala Gly Ile Gly Thr Gly Thr Thr Ala Leu Ile Lys Thr Gln Gln

TTT GAG CAG CTT CAT GCC GCT ATC CAG ACA GAC CTC AAC GAA GTC GAA AAG TCA ATT ACC   1500
Phe Glu Gln Leu His Ala Ala Ile Gln Thr Asp Leu Asn Glu Val Glu Lys Ser Ile Thr

AAC CTA GAA AAG TCA CTG ACC TCG TTG TCT GAA GTA GTC CTA CAG AAC CGC AGA GGC CTA   1560
Asn Leu Glu Lys Ser Leu Thr Ser Leu Ser Glu Val Val Leu Gln Asn Arg Arg Gly Leu

GAT TTG CTA TTC CTA AAG GAG GGA GGT CTC TGC GCA GCC CTA AAA GAA GAA TGT TGT TTT   1620
Asp Leu Leu Phe Leu Lys Glu Gly Gly Leu Cys Ala Ala Leu Lys Glu Glu Cys Cys Phe

TAT GCA GAC CAC ACG GGG CTA GTG AGA GAC AGC ATG GCC AAA TTA AGA GAA AGG CTT AAT   1680
Tyr Ala Asp His Thr Gly Leu Val Arg Asp Ser Met Ala Lys Leu Arg Glu Arg Leu Asn

CAG AGA CAA AAA CTA TTT GAG ACA GGC CAA GGA TGG TTC GAA GGG CTG TTT AAT AGA TCC   1740
Gln Arg Gln Lys Leu Phe Glu Thr Gly Gln Gly Trp Phe Glu Gly Leu Phe Asn Arg Ser

CCC TGG TTT ACC ACC TTA ATC TCC ACC ATC ATG GGA CCT CTA ATA GTA CTC TTA CTG ATC   1800
Pro Trp Phe Thr Thr Leu Ile Ser Thr Ile Met Gly Pro Leu Ile Val Leu Leu Leu Ile

TTA CTC TTT GGA CCT TGC ATT CTC AAT CGA TTG GTC CAA TTT GTT AAA GAC AGG ATC TCA   1860
Leu Leu Phe Gly Pro Cys Ile Leu Asn Arg Leu Val Gln Phe Val Lys Asp Arg Ile Ser

GTG GTC CAG GCT CTG                                                               1875
Val Val Gln Ala Leu
```

FIG. 22 (continued)

CMO Sequence

10 20 30 40 50 60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890

ATG GCG CGT TCA ACG CTC TCA AAA CCC CCT CAA GAT AAG ATT AAC CCG TGG AAG CCC TTA60
Met Ala Arg Ser Thr Leu Ser Lys Pro Pro Gln Asp Lys Ile Asn Pro Trp Lys Pro Leu

ATA GTC ATG GGA GTC CTG TTA GGA GTA GGG ATG GCA GAG AGC CCC CAT CAG GTC TTT AAT20
Ile Val Met Gly Val Leu Leu Gly Val Gly Met Ala Glu Ser Pro His Gln Val Phe Asn

GTA ACC TGG AGA GTC ACC AAC CTG ATG ACT GGG CGT ACC GCC AAT GCC ACC TCC CTC CTG80
Val Thr Trp Arg Val Thr Asn Leu Met Thr Gly Arg Thr Ala Asn Ala Thr Ser Leu Leu

GGA ACT GTA CAA GAT GCC TTC CCA AAA TTA TAT TTT GAT CTA TGT GAT CTG GTC GGA GAG40
Gly Thr Val Gln Asp Ala Phe Pro Lys Leu Tyr Phe Asp Leu Cys Asp Leu Val Gly Glu

GAG TGG GAC CCT TCA GAC CAG GAA CCG TAT GTC GGG TAT GGC TGC AAG TAC CCC GCA GGG00
Glu Trp Asp Pro Ser Asp Gln Glu Pro Tyr Val Gly Tyr Gly Cys Lys Tyr Pro Ala Gly

AGA CAG CGG ACC CGG ACT TTT GAC TTT TAC GTG TGC CCT GGG CAT ACC GTA AAG TCG GGG60
Arg Gln Arg Thr Arg Thr Phe Asp Phe Tyr Val Cys Pro Gly His Thr Val Lys Ser Gly

TGT GGG GGA CCA GGA GAG GGC TAC TGT GGT AAA TGG GGG TGT GAA ACC ACC GGA CAG GCT20
Cys Gly Gly Pro Gly Glu Gly Tyr Cys Gly Lys Trp Gly Cys Glu Thr Thr Gly Gln Ala

TAC TGG AAG CCC ACA TCA TCG TGG GAC CTA ATC TCC CTT AAG CGC GGT AAC ACC CCC TGG80
Tyr Trp Lys Pro Thr Ser Ser Trp Asp Leu Ile Ser Leu Lys Arg Gly Asn Thr Pro Trp

GAC ACG GGA TGC TCT AAA GTT GCC TGT GGC CCC TGC TAC GAC CTC TCC AAA GTA TCC AAT40
Asp Thr Gly Cys Ser Lys Val Ala Cys Gly Pro Cys Tyr Asp Leu Ser Lys Val Ser Asn

TCC TTC CAA GGG GCT ACT CGA GGG GGC AGA TGC AAC CCT CTA GTC CTA GAA TTC ACT GAT00
Ser Phe Gln Gly Ala Thr Arg Gly Gly Arg Cys Asn Pro Leu Val Leu Glu Phe Thr Asp

GCA GGA AAA AAG GCT AAC TGG GAC GGG CCC AAA TCG TGG GGA CTG AGA CTG TAC CGG ACA60
Ala Gly Lys Lys Ala Asn Trp Asp Gly Pro Lys Ser Trp Gly Leu Arg Leu Tyr Arg Thr

GGA ACA GAT CCT ATT ACC ATG TTC TCC CTG ACC CGG CAG GTC CTT AAT GTG GGA CCC CGA20
Gly Thr Asp Pro Ile Thr Met Phe Ser Leu Thr Arg Gln Val Leu Asn Val Gly Pro Arg

GTC CCC ATA GGG CCC AAC CCA GTA TTA CCC GAC CAA AGA CTC CCT TCC TCA CCA ATA GAG80
Val Pro Ile Gly Pro Asn Pro Val Leu Pro Asp Gln Arg Leu Pro Ser Ser Pro Ile Glu

ATT GTA CCG GCT CCA CAG CCA CCT AGC CCC CTC AAT ACC AGT TAC CCC CCT TCC ACT ACG40
Ile Val Pro Ala Pro Gln Pro Pro Ser Pro Leu Asn Thr Ser Tyr Pro Pro Ser Thr Thr

AGT ACA CCC TCA ACC TCC CCT ACA AGT CCA AGT GTC CCA CAG CCA CCC CCA GGT ACC GAA00
Ser Thr Pro Ser Thr Ser Pro Thr Ser Pro Ser Val Pro Gln Pro Pro Pro Gly Thr Glu

AAT AGG CTG CTA AAC TTA GTA GAC GGA GCC TAC CAA GCC CTC AAC CTC ACC AGT CCT GAC60
Asn Arg Leu Leu Asn Leu Val Asp Gly Ala Tyr Gln Ala Leu Asn Leu Thr Ser Pro Asp

AAA ACC CAA GAG TGC TGG TTG TGT CTA GTA GCG GGA CCC CCC TAC TAC GAA GGG GTT GC020
Lys Thr Gln Glu Cys Trp Leu Cys Leu Val Ala Gly Pro Pro Tyr Tyr Glu Gly Val Ala

GTC CTG GGT ACC TAC TCC AAC CAT ACC TCT GCT CCA GCC AAC TGC TCC GTG GCC TCC CAA80
Val Leu Gly Thr Tyr Ser Asn His Thr Ser Ala Pro Ala Asn Cys Ser Val Ala Ser Gln

CAC AAG TTG ACC CTG TCC GAA GTG ACC GGA CAG GGA CTC TGC ATA GGA GCA GTT CCC AAA40
His Lys Leu Thr Leu Ser Glu Val Thr Gly Gln Gly Leu Cys Ile Gly Ala Val Pro Lys

ACA CAT CAG GCC CTA TGT AAT ACC ACC CAG ACA AGC AGT CGA GGG TCC TAT TAT CTA GT200
Thr His Gln Ala Leu Cys Asn Thr Thr Gln Thr Ser Ser Arg Gly Ser Tyr Tyr Leu Val

FIG. 23

CMO Sequence

```
            10                  20                  30                  40                  50                  60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
GCC CCT ACA GGT ACC ATG TGG GCT TGT AGT ACC GGG CTT ACT CCA TGC ATC TCC ACC AC1260
Ala Pro Thr Gly Thr Met Trp Ala Cys Ser Thr Gly Leu Thr Pro Cys Ile Ser Thr Thr

ATA CTG AAC CTT ACC ACT GAT TAT TGT GTT CTT GTC GAA CTC TGG CCA AGA GTC ACC TA1320
Ile Leu Asn Leu Thr Thr Asp Tyr Cys Val Leu Val Glu Leu Trp Pro Arg Val Thr Tyr

CAT TCC CCC AGC TAT GTT TAC GGC CTG TTT GAG AGA TCC AAC CGA CAC AAA AGA GAA CC1380
His Ser Pro Ser Tyr Val Tyr Gly Leu Phe Glu Arg Ser Asn Arg His Lys Arg Glu Pro

GTG TCG TTA ACC CTG GCC CTA TTA TTG GGT GGA CTA ACC ATG GGA GGG ATT GCA GCT GG1440
Val Ser Leu Thr Leu Ala Leu Leu Leu Gly Gly Leu Thr Met Gly Gly Ile Ala Ala Gly

ATA GGG ACG GGG ACC ACT GCC TTA ATT AAA ACC CAG CAG TTT GAG CAG CTT CAT GCC GC1500
Ile Gly Thr Gly Thr Thr Ala Leu Ile Lys Thr Gln Gln Phe Glu Gln Leu His Ala Ala

ATC CAG ACA GAC CTC AAC GAA GTC GAA AAG TCA ATT ACC AAC CTA GAA AAG TCA CTG AC1560
Ile Gln Thr Asp Leu Asn Glu Val Glu Lys Ser Ile Thr Asn Leu Glu Lys Ser Leu Thr

TCG TTG TCT GAA GTA GTC CTA CAG AAC CGC AGA GGC CTA GAT TTG CTA TTC CTA AAG GA1620
Ser Leu Ser Glu Val Val Leu Gln Asn Arg Arg Gly Leu Asp Leu Leu Phe Leu Lys Glu

GGA GGT CTC TGC GCA GCC CTA AAA GAA GAA TGT TGT TTT TAT GCA GAC CAC ACG GGG CT1680
Gly Gly Leu Cys Ala Ala Leu Lys Glu Glu Cys Cys Phe Tyr Ala Asp His Thr Gly Leu

GTG AGA GAC AGC ATG GCC AAA TTA AGA GAA AGG CTT AAT CAG AGA CAA AAA CTA TTT GA1740
Val Arg Asp Ser Met Ala Lys Leu Arg Glu Arg Leu Asn Gln Arg Gln Lys Leu Phe Glu

ACA GGC CAA GGA TGG TTC GAA GGG CTG TTT AAT AGA TCC CCC TGG TTT ACC ACC TTA AT1800
Thr Gly Gln Gly Trp Phe Glu Gly Leu Phe Asn Arg Ser Pro Trp Phe Thr Thr Leu Ile

TCC ACC ATC ATG GGA CCT CTA ATA GTA CTC TTA CTG ATC TTA CTC TTT GGA CCT TGC AT1860
Ser Thr Ile Met Gly Pro Leu Ile Val Leu Leu Leu Ile Leu Leu Phe Gly Pro Cys Ile

CTC AAT CGA TTA GTT CAA TTT GTT AAA GAC AGG ATC TCA GTA GTC CAG GCT TTA GTC CT1920
Leu Asn Arg Leu Val Gln Phe Val Lys Asp Arg Ile Ser Val Val Gln Ala Leu Val Leu

ACT CAA CAA TAC CAC CAG CTA AAA CCA CTA GAA TAC GAG CCA                      1962
Thr Gln Gln Tyr His Gln Leu Lys Pro Leu Glu Tyr Glu Pro
```

FIG. 23 (continued)

CMOdelH Sequence

```
                10          20          30          40          50          60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
ATG GCG CGT TCA ACG CTC TCA AAA CCC CCT CAA GAT AAG ATT AAC CCG TGG AAG CCC TTA60
Met Ala Arg Ser Thr Leu Ser Lys Pro Pro Gln Asp Lys Ile Asn Pro Trp Lys Pro Leu

ATA GTC ATG GGA GTC CTG TTA GGA GTA GGG ATG GCA GAG AGC CCC CAG GTC TTT AAT GTA20
Ile Val Met Gly Val Leu Leu Gly Val Gly Met Ala Glu Ser Pro Gln Val Phe Asn Val

ACC TGG AGA GTC ACC AAC CTG ATG ACT GGG CGT ACC GCC AAT GCC ACC TCC CTC CTG GGA80
Thr Trp Arg Val Thr Asn Leu Met Thr Gly Arg Thr Ala Asn Ala Thr Ser Leu Leu Gly

ACT GTA CAA GAT GCC TTC CCA AAA TTA TAT TTT GAT CTA TGT GAT CTG GTC GGA GAG GAG40
Thr Val Gln Asp Ala Phe Pro Lys Leu Tyr Phe Asp Leu Cys Asp Leu Val Gly Glu Glu

TGG GAC CCT TCA GAC CAG GAA CCG TAT GTC GGG TAT GGC TGC AAG TAC CCC GCA GGG AGA00
Trp Asp Pro Ser Asp Gln Glu Pro Tyr Val Gly Tyr Gly Cys Lys Tyr Pro Ala Gly Arg

CAG CGG ACC CGG ACT TTT GAC TTT TAC GTG TGC CCT GGG CAT ACC GTA AAG TCG GGG TGT60
Gln Arg Thr Arg Thr Phe Asp Phe Tyr Val Cys Pro Gly His Thr Val Lys Ser Gly Cys

GGG GGA CCA GGA GAG GGC TAC TGT GGT AAA TGG GGG TGT GAA ACC ACC GGA CAG GCT TAT20
Gly Gly Pro Gly Glu Gly Tyr Cys Gly Lys Trp Gly Cys Glu Thr Thr Gly Gln Ala Tyr

TGG AAG CCC ACA TCA TCG TGG GAC CTA ATC TCC CTT AAG CGC GGT AAC ACC CCC TGG GAT80
Trp Lys Pro Thr Ser Ser Trp Asp Leu Ile Ser Leu Lys Arg Gly Asn Thr Pro Trp Asp

ACG GGA TGC TCT AAA GTT GCC TGT GGC CCC TGC TAC GAC CTC TCC AAA GTA TCC AAT TCC40
Thr Gly Cys Ser Lys Val Ala Cys Gly Pro Cys Tyr Asp Leu Ser Lys Val Ser Asn Ser

TTC CAA GGG GCT ACT CGA GGG GGC AGA TGC AAC CCT CTA GTC CTA GAA TTC ACT GAT GCA00
Phe Gln Gly Ala Thr Arg Gly Gly Arg Cys Asn Pro Leu Val Leu Glu Phe Thr Asp Ala

GGA AAA AAG GCT AAC TGG GAC GGG CCC AAA TCG TGG GGA CTG AGA CTG TAC CGG ACA GGA60
Gly Lys Lys Ala Asn Trp Asp Gly Pro Lys Ser Trp Gly Leu Arg Leu Tyr Arg Thr Gly

ACA GAT CCT ATT ACC ATG TTC TCC CTG ACC CGG CAG GTC CTT AAT GTG GGA CCC CGA GTC20
Thr Asp Pro Ile Thr Met Phe Ser Leu Thr Arg Gln Val Leu Asn Val Gly Pro Arg Val

CCC ATA GGG CCC AAC CCA GTA TTA CCC GAC CAA AGA CTC CCT TCC TCA CCA ATA GAG ATT80
Pro Ile Gly Pro Asn Pro Val Leu Pro Asp Gln Arg Leu Pro Ser Ser Pro Ile Glu Ile

GTA CCG GCT CCA CAG CCA CCT AGC CCC CTC AAT ACC AGT TAC CCC CCT TCC ACT ACC AGT40
Val Pro Ala Pro Gln Pro Pro Ser Pro Leu Asn Thr Ser Tyr Pro Pro Ser Thr Thr Ser

ACA CCC TCA ACC TCC CCT ACA AGT CCA AGT GTC CCA CAG CCA CCC CCA GGT ACC GAA AAT00
Thr Pro Ser Thr Ser Pro Thr Ser Pro Ser Val Pro Gln Pro Pro Pro Gly Thr Glu Asn

AGG CTG CTA AAC TTA GTA GAC GGA GCC TAC CAA GCC CTC AAC CTC ACC AGT CCT GAC AAA60
Arg Leu Leu Asn Leu Val Asp Gly Ala Tyr Gln Ala Leu Asn Leu Thr Ser Pro Asp Lys

ACC CAA GAG TGC TGG TTG TGT CTA GTA GCG GGA CCC CCC TAC TAC GAA GGG GTT GCC GTG020
Thr Gln Glu Cys Trp Leu Cys Leu Val Ala Gly Pro Pro Tyr Tyr Glu Gly Val Ala Val

CTG GGT ACC TAC TCC AAC CAT ACC TCT GCT CCA GCC AAC TGC TCC GTG GCC TCC CAA CAC080
Leu Gly Thr Tyr Ser Asn His Thr Ser Ala Pro Ala Asn Cys Ser Val Ala Ser Gln His

AAG TTG ACC CTG TCC GAA GTG ACC GGA CAG GGA CTC TGC ATA GGA GCA GTT CCC AAA ACA40
Lys Leu Thr Leu Ser Glu Val Thr Gly Gln Gly Leu Cys Ile Gly Ala Val Pro Lys Thr

CAT CAG GCC CTA TGT AAT ACC ACC CAG ACA AGC AGT CGA GGG TCC TAT TAT CTA GTT GCC200
His Gln Ala Leu Cys Asn Thr Thr Gln Thr Ser Ser Arg Gly Ser Tyr Tyr Leu Val Ala
```

FIG. 24

CMOdelH Sequence

```
                 10                  20                  30                  40                  50                  60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
CCT ACA GGT ACC ATG TGG GCT TGT AGT ACC GGG CTT ACT CCA TGC ATC TCC ACC ACC AT1260
Pro Thr Gly Thr Met Trp Ala Cys Ser Thr Gly Leu Thr Pro Cys Ile Ser Thr Thr Ile

CTG AAC CTT ACC ACT GAT TAT TGT GTT CTT GTC GAA CTC TGG CCA AGA GTC ACC TAT CA1320
Leu Asn Leu Thr Thr Asp Tyr Cys Val Leu Val Glu Leu Trp Pro Arg Val Thr Tyr His

TCC CCC AGC TAT GTT TAC GGC CTG TTT GAG AGA TCC AAC CGA CAC AAA AGA GAA CCG GT1380
Ser Pro Ser Tyr Val Tyr Gly Leu Phe Glu Arg Ser Asn Arg His Lys Arg Glu Pro Val

TCG TTA ACC CTG GCC CTA TTA TTG GGT GGA CTA ACC ATG GGA GGG ATT GCA GCT GGA AT1440
Ser Leu Thr Leu Ala Leu Leu Leu Gly Gly Leu Thr Met Gly Gly Ile Ala Ala Gly Ile

GGG ACG GGG ACC ACT GCC TTA ATT AAA ACC CAG CAG TTT GAG CAG CTT CAT GCC GCT AT1500
Gly Thr Gly Thr Thr Ala Leu Ile Lys Thr Gln Gln Phe Gln Gln Leu His Ala Ala Ile

CAG ACA GAC CTC AAC GAA GTC GAA AAG TCA ATT ACC AAC CTA GAA AAG TCA CTG ACC TC1560
Gln Thr Asp Leu Asn Glu Val Glu Lys Ser Ile Thr Asn Leu Glu Lys Ser Leu Thr Ser

TTG TCT GAA GTA GTC CTA CAG AAC CGC AGA GGC CTA GAT TTG CTA TTC CTA AAG GAG GG1620
Leu Ser Glu Val Val Leu Gln Asn Arg Arg Gly Leu Asp Leu Leu Phe Leu Lys Glu Gly

GGT CTC TGC GCA GCC CTA AAA GAA GAA TGT TGT TTT TAT GCA GAC CAC ACG GGG CTA GT1680
Gly Leu Cys Ala Ala Leu Lys Glu Glu Cys Cys Phe Tyr Ala Asp His Thr Gly Leu Val

AGA GAC AGC ATG GCC AAA TTA AGA GAA AGG CTT AAT CAG AGA CAA AAA CTA TTT GAG AC1740
Arg Asp Ser Met Ala Lys Leu Arg Glu Arg Leu Asn Gln Arg Gln Lys Leu Phe Glu Thr

GGC CAA GGA TGG TTC GAA GGG CTG TTT AAT AGA TCC CCC TGG TTT ACC ACC TTA ATC TC1800
Gly Gln Gly Trp Phe Glu Gly Leu Phe Asn Arg Ser Pro Trp Phe Thr Thr Leu Ile Ser

ACC ATC ATG GGA CCT CTA ATA GTA CTC TTA CTG ATC TTA CTC TTT GGA CCT TGC ATT CT1860
Thr Ile Met Gly Pro Leu Ile Val Leu Leu Leu Ile Leu Leu Phe Gly Pro Cys Ile Leu

AAT CGA TTA GTT CAA TTT GTT AAA GAC AGG ATC TCA GTA GTC CAG GCT TTA GTC CTG AC1920
Asn Arg Leu Val Gln Phe Val Lys Asp Arg Ile Ser Val Val Gln Ala Leu Val Leu Thr

CAA CAA TAC CAC CAG CTA AAA CCA CTA GAA TAC GAG CCA                         1959
Gln Gln Tyr His Gln Leu Lys Pro Leu Glu Tyr Glu Pro
```

FIG. 24 (continued)

TMMO Sequence

```
                 10                  20                  30                  40                  50                  60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
ATG GCG CGT TCA ACG CTC TCA AAA CCC CCT CAA GAT AAG ATT AAC CCG TGG AAG CCC TTA60
Met Ala Arg Ser Thr Leu Ser Lys Pro Pro Gln Asp Lys Ile Asn Pro Trp Lys Pro Leu

ATA GTC ATG GGA GTC CTG TTA GGA GTA GGG ATG GCA GAG AGC CCC CAT CAG GTC TTT AAT20
Ile Val Met Gly Val Leu Leu Gly Val Gly Met Ala Glu Ser Pro His Gln Val Phe Asn

GTA ACC TGG AGA GTC ACC AAC CTG ATG ACT GGG CGT ACC GCC AAT GCC ACC TCG CTC CTG80
Val Thr Trp Arg Val Thr Asn Leu Met Thr Gly Arg Thr Ala Asp Ala Thr Ser Leu Leu

GGA ACT GTA CAA GAT GCC TTC CCA AAA TTA TAT TTT GAT CTA TGT GAT CTG GTC GGA GAG40
Gly Thr Val Gln Asp Ala Phe Pro Lys Leu Tyr Phe Asp Leu Cys Asp Leu Val Gly Glu

GAG TGG GAC CCT TCA GAC CAG GAA CCG TAT GTC GGG TAT GGC TGC AAG TAC CCC GCA GGG00
Glu Trp Asp Pro Ser Asp Gln Glu Pro Tyr Val Gly Tyr Gly Cys Lys Tyr Pro Ala Gly

AGA CAG CGG ACC CGG ACT TTT GAC TTT TAC GTG TGC CCT GGG CAT ACC GTA AAG TCG GGG60
Arg Gln Arg Thr Arg Thr Phe Asp Phe Tyr Val Cys Pro Gly His Thr Val Lys Ser Gly

TGT GGG GGA CCA GGA GAG GGC TAC TGT GGT AAA TGG GGG TGT GAA ACC ACC GGA CAG GCT20
Cys Gly Gly Pro Gly Glu Gly Tyr Cys Gly Lys Trp Gly Cys Glu Thr Thr Gly Gln Ala

TAC TGG AAG CCC ACA TCA TCG TGG GAC CTA ATC TCC CTT AAG CGC GGT AAC ACC CCC TGG80
Tyr Trp Lys Pro Thr Ser Ser Trp Asp Leu Ile Ser Leu Lys Arg Gly Asn Thr Pro Trp

GAC ACG GGA TGC TCT AAA GTT GCC TGT GGC CCC TGC TAC GAC CTC TCC AAA GTA TCC AAT40
Asp Thr Gly Cys Ser Lys Val Ala Cys Gly Pro Cys Tyr Asp Leu Ser Lys Val Ser Asn

TCC TTC CAA GGG GCT ACT CGA GGG GGC AGA TGC AAC CCT CTA GTC CTA GAA TTC ACT GAT00
Ser Phe Gln Gly Ala Thr Arg Gly Gly Arg Cys Asn Pro Leu Val Leu Glu Phe Thr Asp

GCA GGA AAA AAG GCT AAC TGG GAC GGG CCC AAA TCG TGG GGA CTG AGA CTG TAC CGG ACG60
Ala Gly Lys Lys Ala Asn Trp Asp Gly Pro Lys Ser Trp Gly Leu Arg Leu Tyr Arg Thr

GGA ACA GAT CCT ATT ACC ATG TTC TCC CTG ACC CGG CAG GTC CTT AAT GTG GGA CCC CGG20
Gly Thr Asp Pro Ile Thr Met Phe Ser Leu Thr Arg Gln Val Leu Asn Val Gly Pro Arg

GTC CCC ATA GGG CCC AAC CCA GTA TTA CCC GAC CAA AGA CTC CCT TCC TCA CCA ATA GAG80
Val Pro Ile Gly Pro Asn Pro Val Leu Pro Asp Gln Arg Leu Pro Ser Ser Pro Ile Glu

ATT GTA CCG GCT CCA CAG CCA CCT AGC CCC CTC AAT ACC AGT TAC CCC CCT TCC ACT ACC40
Ile Val Pro Ala Pro Gln Pro Pro Ser Pro Leu Asn Thr Ser Tyr Pro Pro Ser Thr Thr

AGT ACA CCC TCA ACC TCC CCT ACA AGT CCA AGT GTC CCA CAG CCA CCC CCA GGA ACT GGA00
Ser Thr Pro Ser Thr Ser Pro Thr Ser Pro Ser Val Pro Gln Pro Pro Pro Gly Thr Gly

GAT AGA CTA CTA GCT CTA GTC AAA GGA GCC TAT CAG GCG CTT AAC CTC ACC AAT CCC GAC60
Asp Arg Leu Leu Ala Leu Val Lys Gly Ala Tyr Gln Ala Leu Asn Leu Thr Asn Pro Asp

AAG ACC CAA GAA TGT TGG CTG TGC TTA GTG TCG GGA CCT CCT TAT TAC GAA GGA GTA GCC20
Lys Thr Gln Glu Cys Trp Leu Cys Leu Val Ser Gly Pro Pro Tyr Tyr Glu Gly Val Ala

GTC GTG GGC ACT TAT ACC AAT CAT TCC ACC GCT CCG GCC AAC TGT ACG GCC ACT TCC CAG80
Val Val Gly Thr Tyr Thr Asn His Ser Thr Ala Pro Ala Asn Cys Thr Ala Thr Ser Gln

CAT AAG CTT ACC CTA TCT GAA GTG ACA GGA CAG GGC CTA TGC ATG GGG GCA GTA CCT AAA40
His Lys Leu Thr Leu Ser Glu Val Thr Gly Gln Gly Leu Cys Met Gly Ala Val Pro Lys

ACT CAC CAG GCC TTA TGT AAC ACC ACC CAA AGC GCC GGC TCA GGA TCC TAC TAC CTT GCT00
Thr His Gln Ala Leu Cys Asn Thr Thr Gln Ser Ala Gly Ser Gly Ser Tyr Tyr Leu Ala
```

FIG. 25

TMMO Sequence

```
                10          20          30          40          50          60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
GCA CCC GCC GGA ACA ATG TGG GCT TGC AGC ACT GGA TTG ACT CCC TGC TTG TCC ACC AGG60
Ala Pro Ala Gly Thr Met Trp Ala Cys Ser Thr Gly Leu Thr Pro Cys Leu Ser Thr Thr

GTG CTC AAT CTA ACC ACA GAT TAT TGT GTA TTA GTT GAA CTC TGG CCC AGA GTA ATT TAG20
Val Leu Asn Leu Thr Thr Asp Tyr Cys Val Leu Val Glu Leu Trp Pro Arg Val Ile Tyr

CAC TCC CCC GAT TAT ATG TAT GGT CAG CTT GAA CAG CGT ACC AAA TAT AAA AGA GAG CGA80
His Ser Pro Asp Tyr Met Tyr Gly Gln Leu Glu Gln Arg Thr Lys Tyr Lys Arg Glu Pro

GTA TCA TTG ACC CTG GCC CTT CTA CTA GGA GGA TTA ACC ATG GGG GGA ATT GCC GCT GGA40
Val Ser Leu Thr Leu Ala Leu Leu Leu Gly Gly Leu Thr Met Gly Gly Ile Ala Ala Gly

ATA GGA ACA GGG ACT ACT GCT CTA ATG GCC ACT CAG CAA TTC CAG CAG CTC CAA GCC GG500
Ile Gly Thr Gly Thr Thr Ala Leu Met Ala Thr Gln Gln Phe Gln Gln Leu Gln Ala Ala

GTA CAG GAT GAT CTC AGG GAG GTT |GAA AAA TCA ATC TCT AAC CTA GAA AAG TCT CTC AC560
Val Gln Asp Asp Leu Arg Glu Val |Glu Lys Ser Ile Ser Asn Leu Glu Lys Ser Leu Thr

TCC CTG TCT GAA GTT GTC CTA CAG |AAT CGA AGG GGC CTA GAC TTG TTA TTT CTA AAA GA520
Ser Leu Ser Glu Val Val Leu Gln |Asn Arg Arg Gly Leu Asp Leu Leu Phe Leu Lys Glu

GGA GGG CTG TGT GCT GCT CTA AAA GAA GAA TGT TGC TTC TAT GCG GAC CAC ACA GGA CT580
Gly Gly Leu Cys Ala Ala Leu Lys Glu Glu Cys Cys Phe Tyr Ala Asp His Thr Gly Leu

GTG AGA GAC AGC ATG GCC AAA TTG AGA GAG AGG CTT AAT CAG AGA CAG AAA CTG TTT GAG40
Val Arg Asp Ser Met Ala Lys Leu Arg Glu Arg Leu Asn Gln Arg Gln Lys Leu Phe Glu

TCA ACT CAA GGA TGG TTT GAG GGA CTG TTT AAC AGA TCC CCT TGG TTT ACC ACC TTG AT800
Ser Thr Gln Gly Trp Phe Glu Gly Leu Phe Asn Arg Ser Pro Trp Phe Thr Thr Leu Ile

TCT ACC ATT ATG GGA CCC CTC ATT GTA CTC CTA ATG ATT TTG CTC TTC GGA CCC TGC AT860
Ser Thr Ile Met Gly Pro Leu Ile Val Leu Leu Met Ile Leu Leu Phe Gly Pro Cys Ile

CTT AAT CGA TTA GTT CAA TTT GTT |AAA GAC AGG ATC TCA GTA GTC CAG GCT TTA GTC CT920
Leu Asn Arg Leu Val Gln Phe Val |Lys Asp Arg Ile Ser Val Val Gln Ala Leu Val Leu

ACT CAA CAA TAC CAC CAG CTA AAA CCA CTA GAA TAC GAG CCA                     1962.
Thr Gln Gln Tyr His Gln Leu Lys Pro Leu Glu Tyr Glu Pro
```

FIG. 25 (continued)

TMMOdelH Sequence

```
                10                  20        |       30                  40                  50                  60
123 456 789 012 345 678 901 234 |567 890 123 456 789 012 345 678 901 234 567 890
ATG GCG CGT TCA ACG CTC TCA AAA |CCC CCT CAA GAT AAG ATT AAC CCG TGG AAG CCC TTA60
Met Ala Arg Ser Thr Leu Ser Lys |Pro Pro Gln Asp Lys Ile Asn Pro Trp Lys Pro Leu

ATA GTC ATG GGA GTC CTG TTA GGA GTA GGG ATG GCA GAG AGC CCC CAG GTC TTT AAT GTA120
Ile Val Met Gly Val Leu Leu Gly Val Gly Met Ala Glu Ser Pro Gln Val Phe Asn Val

ACC TGG AGA GTC ACC AAC CTG ATG ACT GGG CGT ACC GCC AAT GCC ACC TCC CTC CTG GGA180
Thr Trp Arg Val Thr Asn Leu Met Thr Gly Arg Thr Ala Asn Ala Thr Ser Leu Leu Gly

ACT GTA CAA GAT GCC TTC CCA AAA TTA TAT TTT GAT CTA TGT GAT CTG GTC GGA GAG GAG240
Thr Val Gln Asp Ala Phe Pro Lys Leu Tyr Phe Asp Leu Cys Asp Leu Val Gly Glu Glu

TGG GAC CCT TCA GAC CAG GAA CCG TAT GTC GGG TAT GGC TGC AAG TAC CCC GCA GGG AGA300
Trp Asp Pro Ser Asp Gln Glu Pro Tyr Val Gly Tyr Gly Cys Lys Tyr Pro Ala Gly Arg

CAG CGG ACC CGG ACT TTT GAC TTT TAC GTG TGC CCT GGG CAT ACC GTA AAG TCG GGG TGT360
Gln Arg Thr Arg Thr Phe Asp Phe Tyr Val Cys Pro Gly His Thr Val Lys Ser Gly Cys

GGG GGA CCA GGA GAG GGC TAC TGT GGT AAA TGG GGG TGT GAA ACC ACC GGA CAG GCT TAC420
Gly Gly Pro Gly Glu Gly Tyr Cys Gly Lys Trp Gly Cys Glu Thr Thr Gly Gln Ala Tyr

TGG AAG CCC ACA TCA TCG TGG GAC |CTA ATC TCC CTT AAG CGC GGT AAC ACC CCC TGG GAC480
Trp Lys Pro Thr Ser Ser Trp Asp |Leu Ile Ser Leu Lys Arg Gly Asn Thr Pro Trp Asp

ACG GGA TGC TCT AAA GTT GCC TGT |GGC CCC TGC TAC GAC CTC TCC AAA GTA TCC AAT TCC540
Thr Gly Cys Ser Lys Val Ala Cys |Gly Pro Cys Tyr Asp Leu Ser Lys Val Ser Asn Ser

TTC CAA GGG GCT ACT CGA GGG GGC AGA TGC AAC CCT CTA GTC CTA GAA TTC ACT GAT GCA600
Phe Gln Gly Ala Thr Arg Gly Gly Arg Cys Asn Pro Leu Val Leu Glu Phe Thr Asp Ala

GGA AAA AAG GCT AAC TGG GAC GGG CCC AAA TCG TGG GGA CTG AGA CTG TAC CGG ACA GGA660
Gly Lys Lys Ala Asn Trp Asp Gly Pro Lys Ser Trp Gly Leu Arg Leu Tyr Arg Thr Gly

ACA GAT CCT ATT ACC ATG TTC TCC CTG ACC CGG CAG GTC CTT AAT GTG GGA CCC CGA GTC720
Thr Asp Pro Ile Thr Met Phe Ser Leu Thr Arg Gln Val Leu Asn Val Gly Pro Arg Val

CCC ATA GGG CCC AAC CCA GTA TTA CCC GAC CAA AGA CTC CCT TCC TCA CCA ATA GAG ATT780
Pro Ile Gly Pro Asn Pro Val Leu Pro Asp Gln Arg Leu Pro Ser Ser Pro Ile Glu Ile

GTA CCG GCT CCA CAG CCA CCT AGC CCC CTC AAT ACC AGT TAC CCC CCT TCC ACT ACC AGT840
Val Pro Ala Pro Gln Pro Pro Ser Pro Leu Asn Thr Ser Tyr Pro Pro Ser Thr Thr Ser

ACA CCC TCA ACC TCC CCT ACA AGT CCA AGT GTC CCA CAG CCA CCC CCA GGA ACT GGA GAT900
Thr Pro Ser Thr Ser Pro Thr Ser Pro Ser Val Pro Gln Pro Pro Pro Gly Thr Gly Asp

AGA CTA CTA GCT CTA GTC AAA GGA |GCC TAT CAG GCG CTT AAC CTC ACC AAT CCC GAC AAG960
Arg Leu Leu Ala Leu Val Lys Gly |Ala Tyr Gln Ala Leu Asn Leu Thr Asn Pro Asp Lys

ACC CAA GAA TGT TGG CTG TGC TTA |GTG TCG GGA CCT CCT TAT TAC GAA GGA GTA GCG GTT1020
Thr Gln Glu Cys Trp Leu Cys Leu |Val Ser Gly Pro Pro Tyr Tyr Glu Gly Val Ala Val

GTG GGC ACT TAT ACC AAT CAT TCC ACC GCT CCG GCC AAC TGT ACG GCC ACT TCC CAA CAT1080
Val Gly Thr Tyr Thr Asn His Ser Thr Ala Pro Ala Asn Cys Thr Ala Thr Ser Gln His

AAG CTT ACC CTA TCT GAA GTG ACA GGA CAG GGC CTA TGC ATG GGG GCA GTA CCT AAA ACT1140
Lys Leu Thr Leu Ser Glu Val Thr Gly Gln Gly Leu Cys Met Gly Ala Val Pro Lys Thr

CAC CAG GCC TTA TGT AAC ACC ACC CAA AGC GCC GGC TCA GGA TCC TAC TAC CTT GCA GCA1200
His Gln Ala Leu Cys Asn Thr Thr Gln Ser Ala Gly Ser Gly Ser Tyr Tyr Leu Ala Ala
```

FIG. 26

TMMOdelH Sequence

```
                10          20          30          40          50          60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
CCC GCC GGA ACA ATG TGG GCT TGC AGC ACT GGA TTG ACT CCC TGC TTG TCC ACC ACG GTG60
Pro Ala Gly Thr Met Trp Ala Cys Ser Thr Gly Leu Thr Pro Cys Leu Ser Thr Thr Val

CTC AAT CTA ACC ACA GAT TAT TGT GTA TTA GTT GAA CTC TGG CCC AGA GTA ATT TAC CAG20
Leu Asn Leu Thr Thr Asp Tyr Cys Val Leu Val Glu Leu Trp Pro Arg Val Ile Tyr His

TCC CCC GAT TAT ATG TAT GGT CAG CTT GAA CAG CGT ACC AAA TAT AAA AGA GAG CCA GT380
Ser Pro Asp Tyr Met Tyr Gly Gln Leu Glu Gln Arg Thr Lys Tyr Lys Arg Glu Pro Val

TCA TTG ACC CTG GCC CTT CTA CTA GGA GGA TTA ACC ATG GGG GGA ATT GCC GCT GGA AT440
Ser Leu Thr Leu Ala Leu Leu Leu Gly Gly Leu Thr Met Gly Gly Ile Ala Ala Gly Ile

GGA ACA GGG ACT ACT GCT CTA ATG GCC ACT CAG CAA TTC CAG CAG CTC CAA GCC GCA GT500
Gly Thr Gly Thr Thr Ala Leu Met Ala Thr Gln Gln Phe Gln Gln Leu Gln Ala Ala Val

CAG GAT GAT CTC AGG GAG GTT GAA AAA TCA ATC TCT AAC CTA GAA AAG TCT CTC ACT TC560
Gln Asp Asp Leu Arg Glu Val Glu Lys Ser Ile Ser Asn Leu Glu Lys Ser Leu Thr Ser

CTG TCT GAA GTT GTC CTA CAG AAT CGA AGG GGC CTA GAC TTG TTA TTT CTA AAA GAA GG620
Leu Ser Glu Val Val Leu Gln Asn Arg Arg Gly Leu Asp Leu Leu Phe Leu Lys Glu Gly

GGG CTG TGT GCT GCT CTA AAA GAA GAA TGT TGC TTC TAT GCG GAC CAC ACA GGA CTA GT680
Gly Leu Cys Ala Ala Leu Lys Glu Glu Cys Cys Phe Tyr Ala Asp His Thr Gly Leu Val

AGA GAC AGC ATG GCC AAA TTG AGA GAG AGG CTT AAT CAG AGA CAG AAA CTG TTT GAG TC740
Arg Asp Ser Met Ala Lys Leu Arg Glu Arg Leu Asn Gln Arg Gln Lys Leu Phe Glu Ser

ACT CAA GGA TGG TTT GAG GGA CTG TTT AAC AGA TCC CCT TGG TTT ACC ACC TTG ATA TC800
Thr Gln Gly Trp Phe Glu Gly Leu Phe Asn Arg Ser Pro Trp Phe Thr Thr Leu Ile Ser

ACC ATT ATG GGA CCC CTC ATT GTA CTC CTA ATG ATT TTG CTC TTC GGA CCC TGC ATT CT860
Thr Ile Met Gly Pro Leu Ile Val Leu Leu Met Ile Leu Leu Phe Gly Pro Cys Ile Leu

AAT CGA TTA GTT CAA TTT GTT AAA GAC AGG ATC TCA GTA GTC CAG GCT TTA GTC CTG AC920
Asn Arg Leu Val Gln Phe Val Lys Asp Arg Ile Ser Val Val Gln Ala Leu Val Leu Thr

CAA CAA TAC CAC CAG CTA AAA CCA CTA GAA TAC GAG CCA                          1959
Gln Gln Tyr His Gln Leu Lys Pro Leu Glu Tyr Glu Pro
```

FIG. 26 (continued)

CTMMO Sequence

```
                 10                  20                  30                  40                  50                  60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
ATG GCG CGT TCA ACG CTC TCA AAA CCC CCT CAA GAT AAG ATT AAC CCG TGG AAG CCC TTA60
Met Ala Arg Ser Thr Leu Ser Lys Pro Pro Gln Asp Lys Ile Asn Pro Trp Lys Pro Leu

ATA GTC ATG GGA GTC CTG TTA GGA GTA GGG ATG GCA GAG AGC CCC CAT CAG GTC TTT AAT20
Ile Val Met Gly Val Leu Leu Gly Val Gly Met Ala Glu Ser Pro His Gln Val Phe Asn

GTA ACC TGG AGA GTC ACC AAC CTG ATG ACT GGG CGT ACC GCC AAT GCC ACC TCC CTC CTG80
Val Thr Trp Arg Val Thr Asn Leu Met Thr Gly Arg Thr Ala Asn Ala Thr Ser Leu Leu

GGA ACT GTA CAA GAT GCC TTC CCA AAA TTA TAT TTT GAT CTA TGT GAT CTG GTC GGA GAG40
Gly Thr Val Gln Asp Ala Phe Pro Lys Leu Tyr Phe Asp Leu Cys Asp Leu Val Gly Glu

GAG TGG GAC CCT TCA GAC CAG GAA CCG TAT GTC GGG TAT GGC TGC AAG TAC CCC GCA GGG00
Glu Trp Asp Pro Ser Asp Gln Glu Pro Tyr Val Gly Tyr Gly Cys Lys Tyr Pro Ala Gly

AGA CAG CGG ACC CGG ACT TTT GAC TTT TAC GTG TGC CCT GGG CAT ACC GTA AAG TCG GGG60
Arg Gln Arg Thr Arg Thr Phe Asp Phe Tyr Val Cys Pro Gly His Thr Val Lys Ser Gly

TGT GGG GGA CCA GGA GAG GGC TAC TGT GGT AAA TGG GGG TGT GAA ACC ACC GGA CAG GCT20
Cys Gly Gly Pro Gly Glu Gly Tyr Cys Gly Lys Trp Gly Cys Glu Thr Thr Gly Gln Ala

TAC TGG AAG CCC ACA TCA TCG TGG GAC CTA ATC TCC CTT AAG CGC GGT AAC ACC CCC TGG80
Tyr Trp Lys Pro Thr Ser Ser Trp Asp Leu Ile Ser Leu Lys Arg Gly Asn Thr Pro Trp

GAC ACG GGA TGC TCT AAA GTT GCC TGT GGC CCC TGC TAC GAC CTC TCC AAA GTA TCC AAT40
Asp Thr Gly Cys Ser Lys Val Ala Cys Gly Pro Cys Tyr Asp Leu Ser Lys Val Ser Asn

TCC TTC CAA GGG GCT ACT CGA GGG GGC AGA TGC AAC CCT CTA GTC CTA GAA TTC ACT GAT00
Ser Phe Gln Gly Ala Thr Arg Gly Gly Arg Cys Asn Pro Leu Val Leu Glu Phe Thr Asp

GCA GGA AAA AAG GCT AAC TGG GAC GGG CCC AAA TCG TGG GGA CTG AGA CTG TAC CGG ACA60
Ala Gly Lys Lys Ala Asn Trp Asp Gly Pro Lys Ser Trp Gly Leu Arg Leu Tyr Arg Thr

GGA ACA GAT CCT ATT ACC ATG TTC TCC CTG ACC CGG CAG GTC CTT AAT GTG GGA CCC CGA20
Gly Thr Asp Pro Ile Thr Met Phe Ser Leu Thr Arg Gln Val Leu Asn Val Gly Pro Arg

GTC CCC ATA GGG CCC AAC CCA GTA TTA CCC GAC CAA AGA CTC CCT TCC TCA CCA ATA GAG80
Val Pro Ile Gly Pro Asn Pro Val Leu Pro Asp Gln Arg Leu Pro Ser Ser Pro Ile Glu

ATT GTA CCG GCT CCA CAG CCA CCT AGC CCC CTC AAT ACC AGT TAC CCC CCT TCC ACT ACG40
Ile Val Pro Ala Pro Gln Pro Pro Ser Pro Leu Asn Thr Ser Tyr Pro Pro Ser Thr Thr

AGT ACA CCC TCA ACC TCC CCT ACA AGT CCA AGT GTC CCA CAG CCA CCC CCA GGT ACC GAA00
Ser Thr Pro Ser Thr Ser Pro Thr Ser Pro Ser Val Pro Gln Pro Pro Pro Gly Thr Glu

AAT AGG CTG CTA AAC TTA GTA GAC GGA GCC TAC CAA GCC CTC AAC CTC ACC AGT CCT GAG60
Asn Arg Leu Leu Asn Leu Val Asp Gly Ala Tyr Gln Ala Leu Asn Leu Thr Ser Pro Asp

AAA ACC CAA GAG TGC TGG TTG TGT CTA GTA GCG GGA CCC CCC TAC TAC GAA GGG GTT GCC20
Lys Thr Gln Glu Cys Trp Leu Cys Leu Val Ala Gly Pro Pro Tyr Tyr Glu Gly Val Ala

GTC CTG GGT ACC TAC TCC AAC CAT ACC TCT GCT CCA GCC AAC TGC TCC GTG GCC TCC CAG80
Val Leu Gly Thr Tyr Ser Asn His Thr Ser Ala Pro Ala Asn Cys Ser Val Ala Ser Gln

CAC AAG TTG ACC CTG TCC GAA GTG ACC GGA CAG GGA CTC TGC ATA GGA GCA GTT CCC AAA40
His Lys Leu Thr Leu Ser Glu Val Thr Gly Gln Gly Leu Cys Ile Gly Ala Val Pro Lys

ACA CAT CAG GCC CTA TGT AAT ACC ACC CAG ACA AGC AGT CGA GGG TCC TAT TAT CTA GTT00
Thr His Gln Ala Leu Cys Asn Thr Thr Gln Thr Ser Ser Arg Gly Ser Tyr Tyr Leu Val
```

FIG. 27

CTMMO Sequence

```
                10                  20                  30                  40                  50                  60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
GCC CCT ACA GGT ACC ATG TGG GCT TGT AGT ACC GGG CTT ACT CCA TGC ATC TCC ACC AC[illegible]60
Ala Pro Thr Gly Thr Met Trp Ala Cys Ser Thr Gly Leu Thr Pro Cys Ile Ser Thr Thr

ATA CTG AAC CTT ACC ACT GAT TAT TGT GTT CTT GTC GAA CTC TGG CCA AGA GTC ACC TA[illegible]20
Ile Leu Asn Leu Thr Thr Asp Tyr Cys Val Leu Val Glu Leu Trp Pro Arg Val Thr Tyr

CAT TCC CCC AGC TAT GTT TAC GGC CTG TTT GAG AGA TCC AAC CGA CAC AAA AGA GAA CC[illegible]80
His Ser Pro Ser Tyr Val Tyr Gly Leu Phe Glu Arg Ser Asn Arg His Lys Arg Glu Pro

GTG TCG TTA ACC CTG GCC CTA TTA TTG GGT GGA CTA ACC ATG GGG GGA ATT GCC GCT GG[illegible]40
Val Ser Leu Thr Leu Ala Leu Leu Leu Gly Gly Leu Thr Met Gly Gly Ile Ala Ala Gly

ATA GGA ACA GGG ACT ACT GCT CTA ATG GCC ACT CAG CAA TTC CAG CAG CTC CAA GCC GC[illegible]00
Ile Gly Thr Gly Thr Thr Ala Leu Met Ala Thr Gln Gln Phe Gln Gln Leu Gln Ala Ala

GTA CAG GAT GAT CTC AGG GAG GTT GAA AAA TCA ATC TCT AAC CTA GAA AAG TCT CTC AC[illegible]60
Val Gln Asp Asp Leu Arg Glu Val Glu Lys Ser Ile Ser Asn Leu Glu Lys Ser Leu Thr

TCC CTG TCT GAA GTT GTC CTA CAG AAT CGA AGG GGC CTA GAC TTG TTA TTT CTA AAA GA[illegible]20
Ser Leu Ser Glu Val Val Leu Gln Asn Arg Arg Gly Leu Asp Leu Leu Phe Leu Lys Glu

GGA GGG CTG TGT GCT GCT CTA AAA GAA GAA TGT TGC TTC TAT GCG GAC CAC ACA GGA CT[illegible]80
Gly Gly Leu Cys Ala Ala Leu Lys Glu Glu Cys Cys Phe Tyr Ala Asp His Thr Gly Leu

GTG AGA GAC AGC ATG GCC AAA TTG AGA GAG AGG CTT AAT CAG AGA CAG AAA CTG TTT GA[illegible]40
Val Arg Asp Ser Met Ala Lys Leu Arg Glu Arg Leu Asn Gln Arg Gln Lys Leu Phe Glu

TCA ACT CAA GGA TGG TTT GAG GGA CTG TTT AAC AGA TCC CCT TGG TTT ACC ACC TTG AT[illegible]00
Ser Thr Gln Gly Trp Phe Glu Gly Leu Phe Asn Arg Ser Pro Trp Phe Thr Thr Leu Ile

TCT ACC ATT ATG GGA CCC CTC ATT GTA CTC CTA ATG ATT TTG CTC TTC GGA CCC TGC AT[illegible]60
Ser Thr Ile Met Gly Pro Leu Ile Val Leu Leu Met Ile Leu Leu Phe Gly Pro Cys Ile

CTT AAT CGA TTA GTT CAA TTT GTT AAA GAC AGG ATC TCA GTA GTC CAG GCT TTA GTC CT[illegible]20
Leu Asn Arg Leu Val Gln Phe Val Lys Asp Arg Ile Ser Val Val Gln Ala Leu Val Leu

ACT CAA CAA TAC CAC CAG CTA AAA CCA CTA GAA TAC GAG CCA                     1962
Thr Gln Gln Tyr His Gln Leu Lys Pro Leu Glu Tyr Glu Pro
```

FIG. 27 (continued)

CTMMOdelH Sequence

```
                 10                  20                  30                  40                  50                  60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
ATG GCG CGT TCA ACG CTC TCA AAA CCC CCT CAA GAT AAG ATT AAC CCG TGG AAG CCC TTA60
Met Ala Arg Ser Thr Leu Ser Lys Pro Pro Gln Asp Lys Ile Asn Pro Trp Lys Pro Leu

ATA GTC ATG GGA GTC CTG TTA GGA GTA GGG ATG GCA GAG AGC CCC CAG GTC TTT AAT GTA120
Ile Val Met Gly Val Leu Leu Gly Val Gly Met Ala Glu Ser Pro Gln Val Phe Asn Val

ACC TGG AGA GTC ACC AAC CTG ATG ACT GGG CGT ACC GCC AAT GCC ACC TCC CTC CTG GGA180
Thr Trp Arg Val Thr Asn Leu Met Thr Gly Arg Thr Ala Asn Ala Thr Ser Leu Leu Gly

ACT GTA CAA GAT GCC TTC CCA AAA TTA TAT TTT GAT CTA TGT GAT CTG GTC GGA GAG GAG240
Thr Val Gln Asp Ala Phe Pro Lys Leu Tyr Phe Asp Leu Cys Asp Leu Val Gly Glu Glu

TGG GAC CCT TCA GAC CAG GAA CCG TAT GTC GGG TAT GGC TGC AAG TAC CCC GCA GGG AGA300
Trp Asp Pro Ser Asp Gln Glu Pro Tyr Val Gly Tyr Gly Cys Lys Tyr Pro Ala Gly Arg

CAG CGG ACC CGG ACT TTT GAC TTT TAC GTG TGC CCT GGG CAT ACC GTA AAG TCG GGG TGT360
Gln Arg Thr Arg Thr Phe Asp Phe Tyr Val Cys Pro Gly His Thr Val Lys Ser Gly Cys

GGG GGA CCA GGA GAG GGC TAC TGT GGT AAA TGG GGG TGT GAA ACC ACC GGA CAG GCT TAT420
Gly Gly Pro Gly Glu Gly Tyr Cys Gly Lys Trp Gly Cys Glu Thr Thr Gly Gln Ala Tyr

TGG AAG CCC ACA TCA TCG TGG GAC CTA ATC TCC CTT AAG CGC GGT AAC ACC CCC TGG GAT480
Trp Lys Pro Thr Ser Ser Trp Asp Leu Ile Ser Leu Lys Arg Gly Asn Thr Pro Trp Asp

ACG GGA TGC TCT AAA GTT GCC TGT GGC CCC TGC TAC GAC CTC TCC AAA GTA TCC AAT TCC540
Thr Gly Cys Ser Lys Val Ala Cys Gly Pro Cys Tyr Asp Leu Ser Lys Val Ser Asn Ser

TTC CAA GGG GCT ACT CGA GGG GGC AGA TGC AAC CCT CTA GTC CTA GAA TTC ACT GAT GCA600
Phe Gln Gly Ala Thr Arg Gly Gly Arg Cys Asn Pro Leu Val Leu Glu Phe Thr Asp Ala

GGA AAA AAG GCT AAC TGG GAC GGG CCC AAA TCG TGG GGA CTG AGA CTG TAC CGG ACA GGG660
Gly Lys Lys Ala Asn Trp Asp Gly Pro Lys Ser Trp Gly Leu Arg Leu Tyr Arg Thr Gly

ACA GAT CCT ATT ACC ATG TTC TCC CTG ACC CGG CAG GTC CTT AAT GTG GGA CCC CGA GTC720
Thr Asp Pro Ile Thr Met Phe Ser Leu Thr Arg Gln Val Leu Asn Val Gly Pro Arg Val

CCC ATA GGG CCC AAC CCA GTA TTA CCC GAC CAA AGA CTC CCT TCC TCA CCA ATA GAG ATT780
Pro Ile Gly Pro Asn Pro Val Leu Pro Asp Gln Arg Leu Pro Ser Ser Pro Ile Glu Ile

GTA CCG GCT CCA CAG CCA CCT AGC CCC CTC AAT ACC AGT TAC CCC CCT TCC ACT ACC AGT840
Val Pro Ala Pro Gln Pro Pro Ser Pro Leu Asn Thr Ser Tyr Pro Pro Ser Thr Thr Ser

ACA CCC TCA ACC TCC CCT ACA AGT CCA AGT GTC CCA CAG CCA CCC CCA GGT ACC GAA AAT900
Thr Pro Ser Thr Ser Pro Thr Ser Pro Ser Val Pro Gln Pro Pro Pro Gly Thr Glu Asn

AGG CTG CTA AAC TTA GTA GAC GGA GCC TAC CAA GCC CTC AAC CTC ACC AGT CCT GAC AAG960
Arg Leu Leu Asn Leu Val Asp Gly Ala Tyr Gln Ala Leu Asn Leu Thr Ser Pro Asp Lys

ACC CAA GAG TGC TGG TTG TGT CTA GTA GCG GGA CCC CCC TAC TAC GAA GGG GTT GCC GTC1020
Thr Gln Glu Cys Trp Leu Cys Leu Val Ala Gly Pro Pro Tyr Tyr Glu Gly Val Ala Val

CTG GGT ACC TAC TCC AAC CAT ACC TCT GCT CCA GCC AAC TGC TCC GTG GCC TCC CAA CAC1080
Leu Gly Thr Tyr Ser Asn His Thr Ser Ala Pro Ala Asn Cys Ser Val Ala Ser Gln His

AAG TTG ACC CTG TCC GAA GTG ACC GGA CAG GGA CTC TGC ATA GGA GCA GTT CCC AAA ACA1140
Lys Leu Thr Leu Ser Glu Val Thr Gly Gln Gly Leu Cys Ile Gly Ala Val Pro Lys Thr

CAT CAG GCC CTA TGT AAT ACC ACC CAG ACA AGC AGT CGA GGG TCC TAT TAT CTA GTT GCC1200
His Gln Ala Leu Cys Asn Thr Thr Gln Thr Ser Ser Arg Gly Ser Tyr Tyr Leu Val Ala
```

FIG. 28

CTMModelH Sequence

```
                10                  20                  30                  40                  50                  60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
CCT ACA GGT ACC ATG TGG GCT TGT AGT ACC GGG CTT ACT CCA TGC ATC TCC ACC ACC AT1260
Pro Thr Gly Thr Met Trp Ala Cys Ser Thr Gly Leu Thr Pro Cys Ile Ser Thr Thr Ile

CTG AAC CTT ACC ACT GAT TAT TGT GTT CTT GTC GAA CTC TGG CCA AGA GTC ACC TAT CA1320
Leu Asn Leu Thr Thr Asp Tyr Cys Val Leu Val Glu Leu Trp Pro Arg Val Thr Tyr His

TCC CCC AGC TAT GTT TAC GGC CTG TTT GAG AGA TCC AAC CGA CAC AAA AGA GAA CCG GT1380
Ser Pro Ser Tyr Val Tyr Gly Leu Phe Glu Arg Ser Asn Arg His Lys Arg Glu Pro Val

TCG TTA ACC CTG GCC CTA TTA TTG GGT GGA CTA ACC ATG GGG GGA ATT GCC GCT GGA AT1440
Ser Leu Thr Leu Ala Leu Leu Leu Gly Gly Leu Thr Met Gly Gly Ile Ala Ala Gly Ile

GGA ACA GGG ACT ACT GCT CTA ATG GCC ACT CAG CAA TTC CAG CAG CTC CAA GCC GCA GT1500
Gly Thr Gly Thr Thr Ala Leu Met Ala Thr Gln Gln Phe Gln Gln Leu Gln Ala Ala Val

CAG GAT GAT CTC AGG GAG GTT GAA AAA TCA ATC TCT AAC CTA GAA AAG TCT CTC ACT TC1560
Gln Asp Asp Leu Arg Glu Val Glu Lys Ser Ile Ser Asn Leu Glu Lys Ser Leu Thr Ser

CTG TCT GAA GTT GTC CTA CAG AAT CGA AGG GGC CTA GAC TTG TTA TTT CTA AAA GAA GG1620
Leu Ser Glu Val Val Leu Gln Asn Arg Arg Gly Leu Asp Leu Leu Phe Leu Lys Glu Gly

GGG CTG TGT GCT GCT CTA AAA GAA GAA TGT TGC TTC TAT GCG GAC CAC ACA GGA CTA GT1680
Gly Leu Cys Ala Ala Leu Lys Glu Glu Cys Cys Phe Tyr Ala Asp His Thr Gly Leu Val

AGA GAC AGC ATG GCC AAA TTG AGA GAG AGG CTT AAT CAG AGA CAG AAA CTG TTT GAG TC1740
Arg Asp Ser Met Ala Lys Leu Arg Glu Arg Leu Asn Gln Arg Gln Lys Leu Phe Glu Ser

ACT CAA GGA TGG TTT GAG GGA CTG TTT AAC AGA TCC CCT TGG TTT ACC ACC TTG ATA TC1800
Thr Gln Gly Trp Phe Glu Gly Leu Phe Asn Arg Ser Pro Trp Phe Thr Thr Leu Ile Ser

ACC ATT ATG GGA CCC CTC ATT GTA CTC CTA ATG ATT TTG CTC TTC GGA CCC TGC ATT CT1860
Thr Ile Met Gly Pro Leu Ile Val Leu Leu Met Ile Leu Leu Phe Gly Pro Cys Ile Leu

AAT CGA TTA GTT CAA TTT GTT AAA GAC AGG ATC TCA GTA GTC CAG GCT TTA GTC CTG AC1920
Asn Arg Leu Val Gln Phe Val Lys Asp Arg Ile Ser Val Val Gln Ala Leu Val Leu Thr

CAA CAA TAC CAC CAG CTA AAA CCA CTA GAA TAC GAG CCA                        1959
Gln Gln Tyr His Gln Leu Lys Pro Leu Glu Tyr Glu Pro
```

FIG. 28 (continued)

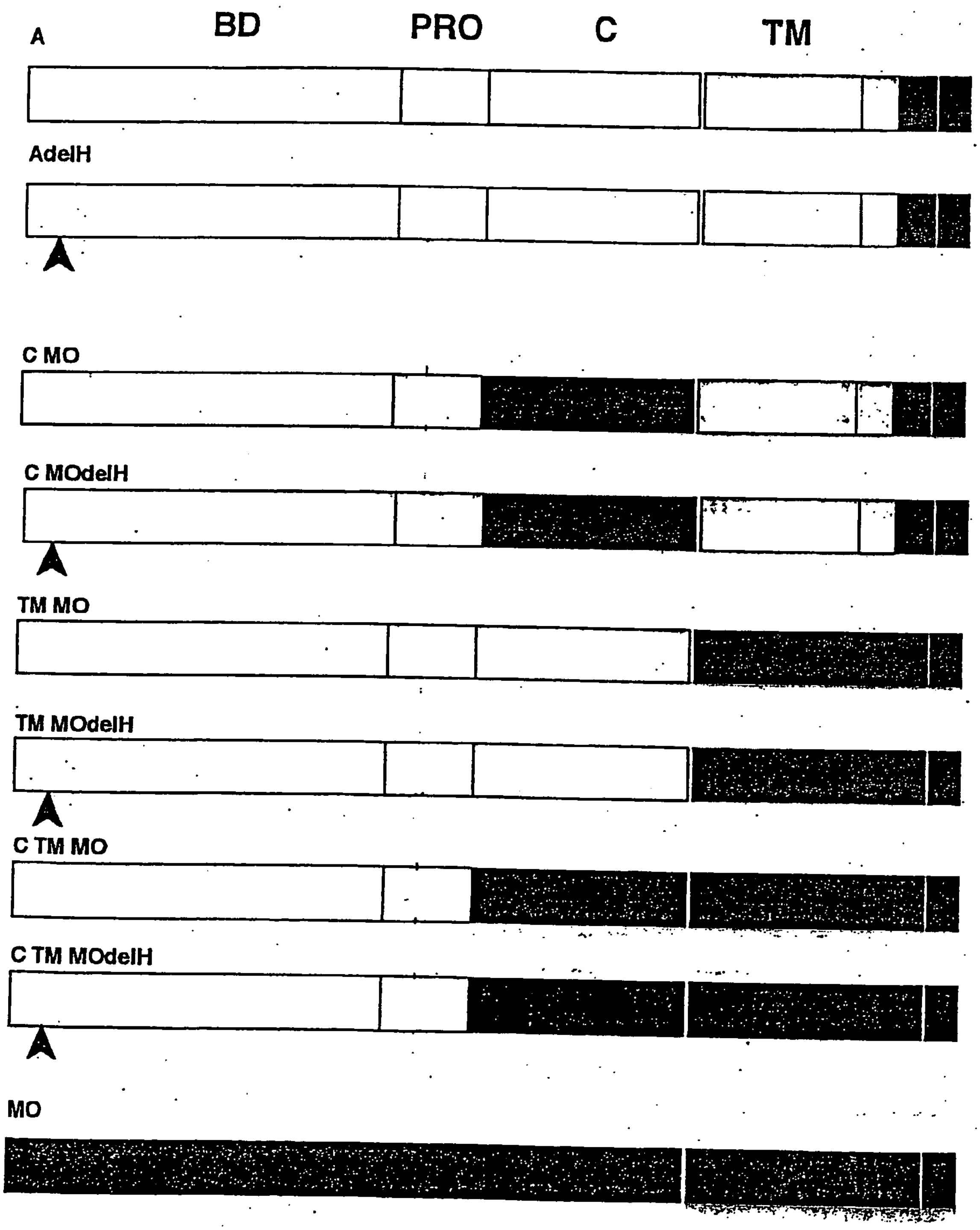
A
BD
PRO
C
TM
AdelH
C MO
C MOdelH
TM MO
TM MOdelH
C TM MO
C TM MOdelH
MO

FIG. 29 A

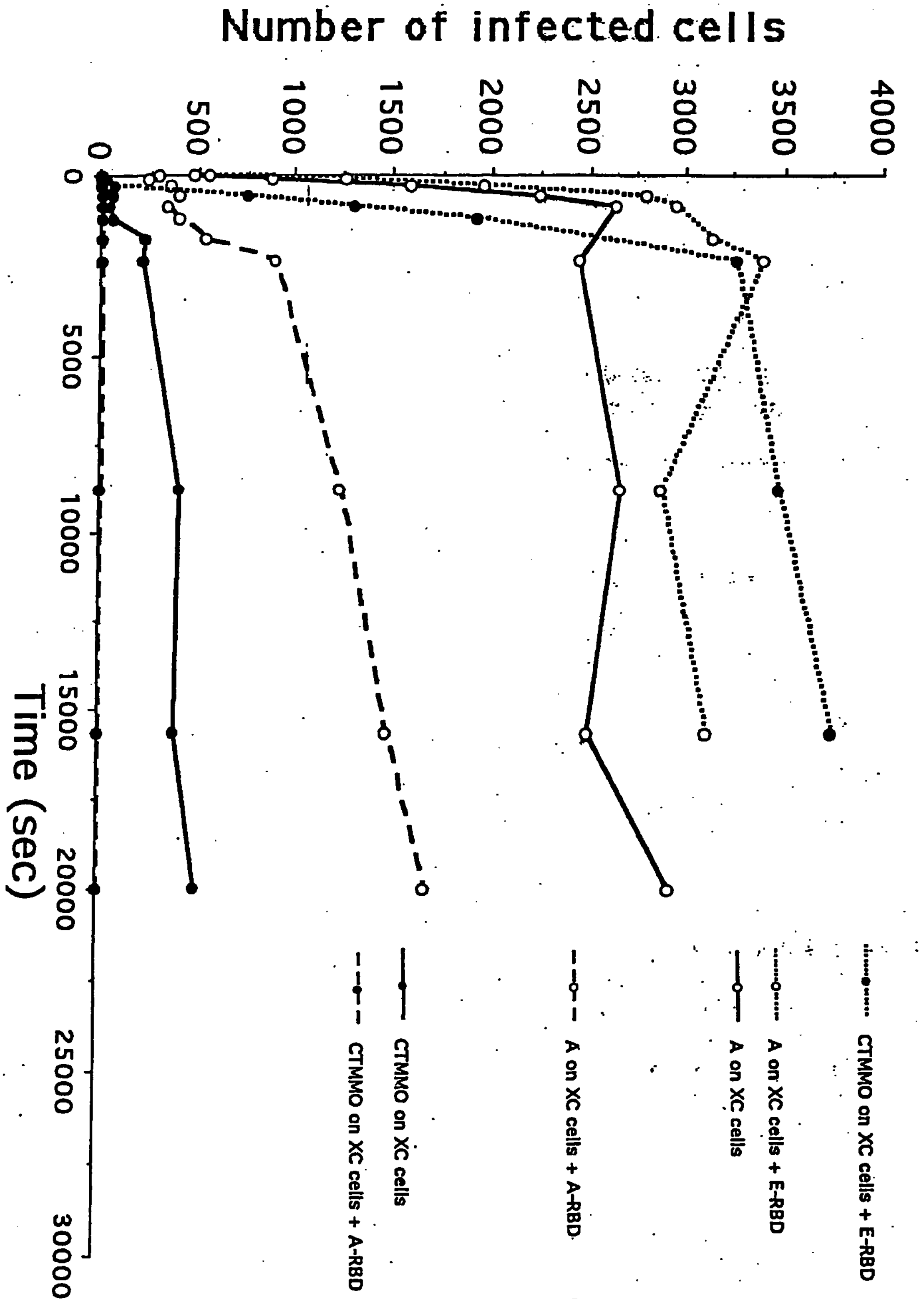
Number of infected cells
0
500
1000
1500
2000
2500
3000
3500
4000
Time (sec)
0
5000
10000
15000
20000
25000
30000
CTMMO on XC cells + E-RBD
A on XC cells + E-RBD
A on XC cells
A on XC cells + A-RBD
CTMMO on XC cells
CTMMO on XC cells + A-RBD

FIG.29 B

34.1-A Sequence

```
                 10          20          30          40          50          60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
ATG GCG CGT TCA ACG CTC TCA AAA CCC CTT AAA AAT AAG GTT AAC CCG CGA GGC CCC CTA60
Met Ala Arg Ser Thr Leu Ser Lys Pro Leu Lys Asn Lys Val Asn Pro Arg Gly Pro Leu

ATC CCC TTA ATT CTT CTG ATG CTC AGA GGG GTC AGT ACT GCG GCC CAG CCG GCC ATG GCC120
Ile Pro Leu Ile Leu Leu Met Leu Arg Gly Val Ser Thr Ala Ala Gln Pro Ala Met Ala

CAG GTT CAG CTT CAG CAG TCT GGA GCT GAA CTG GTA AAG CCT GGG GCT TCA GTG AAG TTG180
Gln Val Gln Leu Gln Gln Ser Gly Ala Glu Leu Val Lys Pro Gly Ala Ser Val Lys Leu

TCC TGC AAG GCT TCT GGC TAC ACC TTC ACA AGC TAT GAT ATA AAC TGG GTG AGG CAG AGG240
Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr Asp Ile Asn Trp Val Arg Gln Arg

CCT GAA CAG GGA CTT GAG TGG ATT GGA TGG ATT TTT CCT GGA GAT GGT AGT ACT AAG TAC300
Pro Glu Gln Gly Leu Glu Trp Ile Gly Trp Ile Phe Pro Gly Asp Gly Ser Thr Lys Tyr

AAT GAG AAG TTC AAG GGC AAG GCC ACA CTG ACT ACA GAC AAA TCC TCC AGC ACA GCC TAC360
Asn Glu Lys Phe Lys Gly Lys Ala Thr Leu Thr Thr Asp Lys Ser Ser Ser Thr Ala Tyr

ATG CAG CTC AGC AGG CTG ACA TCT GAA GAC TCT GCT ATC TAT TTC TGT GCA AGA ATG GCC420
Met Gln Leu Ser Arg Leu Thr Ser Glu Asp Ser Ala Ile Tyr Phe Cys Ala Arg Met Ala

GAC TAT GGT CAC TAC GGC GCT GTG GAC TAC TGG GGT CAA GGG ACC ACG GTC ACC GTC TCA480
Asp Tyr Gly His Tyr Gly Ala Val Asp Tyr Trp Gly Gln Gly Thr Thr Val Thr Val Ser

TCA GGT GGA GGC GGT TCA GGC GGA GGT GGC TCT GGC GGT GGC GGA TCG GAC ATC CAG ATG540
Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Gln Met

ACT CAG TCT CCA TCC TCC CTG TCT GCC TCT CTG GGA GAC AGA GTC ACC ATC AGT TGC AGG600
Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Leu Gly Asp Arg Val Thr Ile Ser Cys Arg

GCA AGT CAG GAC ATT AGC AAT TAT TTA AAC TGG TAT CAG CAG ACA CCA GAT GGA ACT GTT660
Ala Ser Gln Asp Ile Ser Asn Tyr Leu Asn Trp Tyr Gln Gln Thr Pro Asp Gly Thr Val

AAA CTC CTG ATC TAC TAC ACA TCA AGA TTA CAC TCA GGA GTC CCA TCA AGG TTC AGT GGC720
Lys Leu Leu Ile Tyr Tyr Thr Ser Arg Leu His Ser Gly Val Pro Ser Arg Phe Ser Gly

AGT GGG TCT GGA ACA GAT TAT TCT CTC ACC ATT AGC AAC CTG GAG CAA GAA GAT ATT GCC780
Ser Gly Ser Gly Thr Asp Tyr Ser Leu Thr Ile Ser Asn Leu Glu Gln Glu Asp Ile Ala

ACT TAC TTT TGC CAA CAG GGT TAT ACG CTT CCG TGG ACG TTC GGT GGA GGC ACC AAG CTG840
Thr Tyr Phe Cys Gln Gln Gly Tyr Thr Leu Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu

GAG ATC AAA CGG GCG GCC GCA GGA CCC CGA GTC CCC ATA GGG CCC AAC CCA GTA TTA CCG900
Glu Ile Lys Arg Ala Ala Ala Gly Pro Arg Val Pro Ile Gly Pro Asn Pro Val Leu Pro

GAC CAA AGA CTC CCT TCC TCA CCA ATA GAG ATT GTA CCG GCT CCA CAG CCA CCT AGC CCG960
Asp Gln Arg Leu Pro Ser Ser Pro Ile Glu Ile Val Pro Ala Pro Gln Pro Pro Ser Pro

CTC AAT ACC AGT TAC CCC CCT TCC ACT ACC AGT ACA CCC TCA ACC TCC CCT ACA AGT CCA1020
Leu Asn Thr Ser Tyr Pro Pro Ser Thr Thr Ser Thr Pro Ser Thr Ser Pro Thr Ser Pro

AGT GTC CCA CAG CCA CCC CCA GGA ACT GGA GAT AGA CTA CTA GCT CTA GTC AAA GGA GCC1080
Ser Val Pro Gln Pro Pro Pro Gly Thr Gly Asp Arg Leu Leu Ala Leu Val Lys Gly Ala

TAT CAG GCG CTT AAC CTC ACC AAT CCC GAC AAG ACC CAA GAA TGT TGG CTG TGC TTA GTG1140
Tyr Gln Ala Leu Asn Leu Thr Asn Pro Asp Lys Thr Gln Glu Cys Trp Leu Cys Leu Val
```

FIG. 30

34.1-A Sequence

```
                10          20          30          40          50          60
123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
TCG GGA CCT CCT TAT TAC GAA GGA GTA GCG GTC GTG GGC ACT TAT ACC AAT CAT TCC AC1200
Ser Gly Pro Pro Tyr Tyr Glu Gly Val Ala Val Val Gly Thr Tyr Thr Asn His Ser Thr

GCT CCG GCC AAC TGT ACG GCC ACT TCC CAA CAT AAG CTT ACC CTA TCT GAA GTG ACA GG1260
Ala Pro Ala Asn Cys Thr Ala Thr Ser Gln His Lys Leu Thr Leu Ser Glu Val Thr Gly

CAG GGC CTA TGC ATG GGG GCA GTA CCT AAA ACT CAC CAG GCC TTA TGT AAC ACC ACC CA1320
Gln Gly Leu Cys Met Gly Ala Val Pro Lys Thr His Gln Ala Leu Cys Asn Thr Thr Gln

AGC GCC GGC TCA GGA TCC TAC TAC CTT GCA GCA CCC GCC GGA ACA ATG TGG GCT TGC AG1380
Ser Ala Gly Ser Gly Ser Tyr Tyr Leu Ala Ala Pro Ala Gly Thr Met Trp Ala Cys Ser

ACT GGA TTG ACT CCC TGC TTG TCC ACC ACG GTG CTC AAT CTA ACC ACA GAT TAT TGT GT1440
Thr Gly Leu Thr Pro Cys Leu Ser Thr Thr Val Leu Asn Leu Thr Thr Asp Tyr Cys Val

TTA GTT GAA CTC TGG CCC AGA GTA ATT TAC CAC TCC CCC GAT TAT ATG TAT GGT CAG CT1500
Leu Val Glu Leu Trp Pro Arg Val Ile Tyr His Ser Pro Asp Tyr Met Tyr Gly Gln Leu

GAA CAG CGT ACC AAA TAT AAA AGA GAG CCA GTA TCA TTG ACC CTG GCC CTT CTA CTA GG1560
Glu Gln Arg Thr Lys Tyr Lys Arg Glu Pro Val Ser Leu Thr Leu Ala Leu Leu Leu Gly

GGA TTA ACC ATG GGA GGG ATT GCA GCT GGA ATA GGG ACG GGG ACC ACT GCC TTA ATT AA1620
Gly Leu Thr Met Gly Gly Ile Ala Ala Gly Ile Gly Thr Gly Thr Thr Ala Leu Ile Lys

ACC CAG CAG TTT GAG CAG CTT CAT GCC GCT ATC CAG ACA GAC CTC AAC GAA GTC GAA AA1680
Thr Gln Gln Phe Glu Gln Leu His Ala Ala Ile Gln Thr Asp Leu Asn Glu Val Glu Lys

TCA ATT ACC AAC CTA GAA AAG TCA CTG ACC TCG TTG TCT GAA GTA GTC CTA CAG AAC CG1740
Ser Ile Thr Asn Leu Glu Lys Ser Leu Thr Ser Leu Ser Glu Val Val Leu Gln Asn Arg

AGA GGC CTA GAT TTG CTA TTC CTA AAG GAG GGA GGT CTC TGC GCA GCC CTA AAA GAA GA1800
Arg Gly Leu Asp Leu Leu Phe Leu Lys Glu Gly Gly Leu Cys Ala Ala Leu Lys Glu Glu

TGT TGT TTT TAT GCA GAC CAC ACG GGG CTA GTG AGA GAC AGC ATG GCC AAA TTA AGA GA1860
Cys Cys Phe Tyr Ala Asp His Thr Gly Leu Val Arg Asp Ser Met Ala Lys Leu Arg Glu

AGG CTT AAT CAG AGA CAA AAA CTA TTT GAG ACA GGC CAA GGA TGG TTC GAA GGG CTG TT1920
Arg Leu Asn Gln Arg Gln Lys Leu Phe Glu Thr Gly Gln Gly Trp Phe Glu Gly Leu Phe

AAT AGA TCC CCC TGG TTT ACC ACC TTA ATC TCC ACC ATC ATG GGA CCT CTA ATA GTA CT1980
Asn Arg Ser Pro Trp Phe Thr Thr Leu Ile Ser Thr Ile Met Gly Pro Leu Ile Val Leu

TTA CTG ATC TTA CTC TTT GGA CCT TGC ATT CTC AAT CGA TTG GTC CAA TTT GTT AAA GA2040
Leu Leu Ile Leu Leu Phe Gly Pro Cys Ile Leu Asn Arg Leu Val Gln Phe Val Lys Asp

AGG ATC TCA GTG GTC CAG GCT CTG GTT TTG ACT CAG CAA TAT CAC CAG CTA AAA CCC AT2100
Arg Ile Ser Val Val Gln Ala Leu Val Leu Thr Gln Gln Tyr His Gln Leu Lys Pro Ile

GAG TAC GAG CCA TGA                                                         2115
Glu Tyr Glu Pro ...
```

FIG. 30 (continued)